

;; TITLE OF INVENTION: Human GPR4 G-Protein Coupled Receptor and
;; Nucleotides Encoding Same (As Amended)
;; NUMBER OF SEQUENCES: 30
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;; STREET: 1100 NEW YORK AVE., NW, SUITE 600
;; CITY: WASHINGTON
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/024,494
;; FILING DATE: 21-Dec-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/465,973
;; FILING DATE: 06-JUN-1995
;; APPLICATION NUMBER: PCT/US95/04079
;; FILING DATE: 30-MAR-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEFFE, ERIC K.
;; REGISTRATION NUMBER: 36,688
;; REFERENCE/DOCKET NUMBER: 1488.1140001/EKS/HCC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-2600
;; TELEFAX: 202-371-2540
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: No. US20030044898A1 Relevant
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-024-494-30

Query Match 18.4%; Score 397; DB 9; Length 302;
Best Local Similarity 34.6%; Pred. No. 7.6e-25;
Matches 103; Conservative 48; Mismatches 123; Indels 24; Gaps 10;
QY 108 TSSLSTK-----LIPAIYLLVFVGVGPANAVTLMWLF--RTRSICTTVFYTNLAIDF 159
DB 1 SSSCTSDSKYLYGCVFSWVGLIANCVAIYFTFLKVRNE-TTITYMLMLAISDL 59
QY 160 LFCVTLFPFKIAYHLNGNNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHPFT 219
DB 60 LFVFTLPFRI-YVYVVRNWPFGDVLCKISVTLFVYNNYGSILFULTCISVDRFLAIVHPPR 118
QY 220 YRGLPHTYALVTCGLWATVFLYMLPFFFLKQEYLVQPDITTDHVDHNTCESS--SPFQ 278
DB 119 SKTLRTKRNARIYCVAVWITVLAGSTPASFPQSTNRQNNTEQRTCFE--NFPESTWKTYL 176
QY 279 LVYFISLAPFGFLIPFVLIYCYAAIIRTLN---AYDHRWLWVVKASLLI---LVIFTIC 332
DB 177 SRVIFIEIVGFFIPLILNVTCTWVLRUNLKPLTISRNLKSKKKVLMFVHLVIFPCF 236
QY 333 FAFSNIILIIH---HANYYYNNT--DGLYFIYLIACLSLNSCLDPFLYFLMSKTRN 385
DB 237 FVPYNTLIYLSMRQTWNCNSVVTAVRTMYPTVLTICIAVSNCCFPIVYVYFTSDINS 294

RESULT 26

US-10-270-144-4
; Sequence 4, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED

;; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
;; FILE REFERENCE: CL000750CON
;; CURRENT APPLICATION NUMBER: US/10/270,144
;; PRIOR FILING DATE: 2002-10-15
;; PRIOR APPLICATION NUMBER: 60/205,196
;; PRIOR FILING DATE: 2000-05-18
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 299
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-10-270-144-4
Query Match 18.4%; Score 396.5; DB 9; Length 299;
Best Local Similarity 32.5%; Pred. No. 8.3e-25;
Matches 95; Conservative 57; Mismatches 101; Indels 39; Gaps 11;
QY 116 IPAIYLLVFVGVGPANAVTLMWLFRTSRIC-TTVFYTNLAIDFLFCVTLFPFKIAYHLN 174
DB 13 LPNAVILVFIIGFLGNSVAIMWVFMKFPWGSIGSVYMFNLALADFLYVLTLPALIFYFN 72
QY 175 GNNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHPFTYRGLPHTYALVTCG 234
DB 73 KTDWTFGDAMCKLQRFIFHVNLYGSILFULTCISAHRYSGVYVPLKSLGRLKKNAIYVS 132
QY 235 LWNAVTFYMLPFFFLKQEY---LVQPDITTDHVDHNTCESSSPFQLYYFI-----SL 285
DB 133 LVWLIVVVAISPIL-----FYSGTGTRKNKTVTCVDT-----TSDYLSRYFIYSMCTTV 182
QY 286 AFFGLFIPFLVLIYCYAAIIRTL--NAYDH-----RWLWYVKASLLILVIFTICFAPSNI 338
DB 183 AM--FCIPLVLILGCVGLIVKALIVNLDLNSPLRRKSIYLV--IIVLTVFAVSYPHV 237
QY 339 ILIIH-HANYYYN-----TDGLYFIYLIACLSLNSCLDPFLYFLMSKT 383
DB 238 MKTMNLRARLDFQTPMCDNFDRVYATQVTRGLASLNSCVDPILYFLAGDT 289
RESULT 27
US-09-779-679-25
; Sequence 25, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Nuralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sau
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09

Db 183 SVTSCLELN--LYKIAKLOTMYIALV-VGCLLPFTTSLICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLPYPHTLTHTWKVGLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 32
US-09-779-679-26
; Sequence 26, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779, 679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180, 929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-26
Query Match 17.5%; Score 376.5; DB 9; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;
QY 82 TITVKIKCPESASHLVKVNATGKLTSSLSSTKLIPAIYLLVFGVGPANAVTLMLFFR 141
Db 11 SISVSEMEPNGTFSNNRNCIT----ENFKREFPIVYLIIFFWGLGSLSIY-VFLQ 65
QY 142 --TRSICTTFTYNTLAIDFLFCVTLPPKIAIYHLNGNNWVFGVLCRAITVIFYGNMYS 199
Db 66 PYKSTSVNVFMLNLAISDLSTLFPFRADYLRGNSWIFGDLACRIMSISLYVNMYS 125
QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQ 259
Db 126 IYFLTVLSVVRFLAMVHPFRLHVTIRSAILCGIIV--ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDHVDHNTCESSPPQLYYFISLAFGLPIPLFYLLIYCYAAIIRTINAVD----- 312
Db 183 SVTSCLELN--LYKIAKLOTMYIALV-VGCLLPFTTSLICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLPYPHTLTHTWKVGLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 34
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.

Db 126 IYFLTVLSVVRFLAMVHPFRLHVTIRSAILCGIIV--ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDHVDHNTCESSPPQLYYFISLAFGLPIPLFYLLIYCYAAIIRTINAVD----- 312
Db 183 SVTSCLELN--LYKIAKLOTMYIALV-VGCLLPFTTSLICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLPYPHTLTHTWKVGLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 33
US-09-826-791-6
; Sequence 6, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-791-6
Query Match 17.5%; Score 376.5; DB 10; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;
QY 82 TITVKIKCPESASHLVKVNATGKLTSSLSSTKLIPAIYLLVFGVGPANAVTLMLFFR 141
Db 11 SISVSEMEPNGTFSNNRNCIT----ENFKREFPIVYLIIFFWGLGSLSIY-VFLQ 65
QY 142 --TRSICTTFTYNTLAIDFLFCVTLPPKIAIYHLNGNNWVFGVLCRAITVIFYGNMYS 199
Db 66 PYKSTSVNVFMLNLAISDLSTLFPFRADYLRGNSWIFGDLACRIMSISLYVNMYS 125
QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQ 259
Db 126 IYFLTVLSVVRFLAMVHPFRLHVTIRSAILCGIIV--ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDHVDHNTCESSPPQLYYFISLAFGLPIPLFYLLIYCYAAIIRTINAVD----- 312
Db 183 SVTSCLELN--LYKIAKLOTMYIALV-VGCLLPFTTSLICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLPYPHTLTHTWKVGLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 34
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.


```

; TITLE OF INVENTION: No. US20020150901AmeI Nucleic Acids, Polypeptides, Methods of Making
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-7

Query Match          17.5%; Score 376.5; DB 10; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;

QY      82  TITVTKICPRESASHLVKVNATWGLYTSSUSTKLIPAIYLLVFGVGPANATVLMWLFPR 141
Db      82  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      11  SISVSEMEPNGTFSNNNSRNCIT----ENPKREFFPITVYLIIFPWGLNGLSIY-VFLQ 65
Db      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      142 --TRSICTTVFYNNLAIDELFCVTLPFKIAYHLGNWNVFGEVLCRATTVIFYGNMYCS 199
Db      142 --PYKSTSVNVFMNLALISDLLFTSLDFRADYLRGNSWIFGDLACRIMSISLYVNNYSS 125
QY      200  ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFFLKOEYLYVQP 259
Db      200  IYFLTIVSVVRFLAMVHPFLLHVTISRSAWILCGIIW---ILIMASSIMLLDSGSEQNG 182
QY      260  DIITCDVHNTCSSSPFQIYYFISLAFFGLFPFVLLIYCYYAAIRTLNAYD----- 312
Db      260  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      183  SVTSCUELN--LYKIAKIQTMNYIALV-VGCLLPFFFTLSICYLLIIRVLLKVEVPESGLR 239
Db      183  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      313  --HRLWVYWKASLLILVIFTICFAPSNIILIIHHANYYN-NTDGLYFYFLIALCLGSLN 369
Db      313  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      240  VSRHKA--LTIITIIILIFFCLFPHYTLRTVHLTTWKVGLCKDRHLKALVITLAAAN 297
Db      240  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      370  SCLDPFLYF 378
Db      370  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      298  ACFNPLLY 306
Db      298  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 35
US-09-866-230--9
; Sequence 9, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION: Andrew, et al.
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230--9

Query Match          17.5%   Score 376.5; DB 10; Length 346;
Best Local Similarity 30.1%; Pred.No.4.le-23;
Matches      93; Conservative    62; Mismatches 129; Indels     25; Gaps       9;

Qy      82 TITKIKCPESAGHLHWKATMGYLTSLSLKLIPIAYLLVFVVGVPANAVTLMLFFR 141
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      11 SISVSEMEPNCTFSNNNSRNCTI---ENFKREFPPIVYLIIFPWGVLGNGLSIY-VFLQ 65
        :|::||::||::||::||::||::||::||::||::||::||::||::||:

Qy      142 --TRSICTTTFVYTNLAIADLFECVTLPFKIAHYHNGNWNVGEVLCRAATTIVIFYGNMYCS 199
        :|::||::||::||::||::||::||::||::||::||::||::||::||:

```

Db	66	PYKSTSVNVPMNLNALSDDLFI	STLPLPRADYYLRGSNWI	FGDLACRIMSYSLYVNMYS	125
Qy	200	ILLACISINRYLAIVHPFTYRG	LPGKHTYALVTCGLVMATVFL	YMLPFFILKOEYYLVQP	259
Db	126	IYFTLVSVRFLAMVHPFRLLH	VTIRSAMILCGIIW---	ILIMASSIMLLDSGSEQNG	182
Qy	260	DIITCHDVHNTCESSSPQLYF	YSLAFGGFLIPVLIIYCVA	AIITLNAYD-----	312
Db	183	SVTSCLEUN--LYKIAKLQTN	YIALV-VGCLLPFTTLCI	YLLIIRVLLKVEVPESGR	239
Qy	313	--HRMLWYVKASLHILVIFT	ICAPSNIIILIHANYYN	-NTDGLYFYLIYALCIGSLN	369
Db	240	VSRHKA--LTTIIITLIIF	FLCFLPYHTLRTVHLT	TKVGLCKDLHKLAVITL	297
Qy	370	SCLOPFLYF	378		
Db	298	ACFNPLLY	306		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

[illegible]

Db	5	NSSHCFYNDSPKVTLYGCMFMSVFLGLISNCVAIYF-----ICVLKVRNTEITTYMIN	58
Qy	154	LAIADFLFCVTLFPFKIAYHLGNMNVGCVLCRATTVIFYGNMYCSILLACISINRYLA	213
Db	59	LAMSDLLFVFTLPPRI-FVFETRNWPGDGLCKISVMLFVTNMYGSILFUTCISVDRELA	117
Qy	214	IVHPFFYRGLPKHTYALVTCGLWATVFLYMLPFLLKQEYVLVOPDITTCCHDVHNTCS	273
Db	118	IVVPFKSKTLRTYRNAKIVCTGVMLTVGGSA-----AVFQSTHSGNNASACEPE	170
Qy	274	SSPQOLY--YFISLAF-----GFLIPPVLIICYAAIIRTLN---AYDHRWLWYVKASLL	324
Db	171	NPPEATWKYLSRIVFIEIVGFFPIPLINVTCSMWLKTLTTPVTLRSRKINKTKVLKM	230
Qy	325	I-----IWFITFCAPSNIILIH---HANYYNNT--DGLYFIYLIALCGLSNSCLDPFL	376
Db	231	IFVHLIIFCFVPVNNILYSLVRTQTQVNCVSVAARTMTVITILICIAVNSCCFPDIP	290
Qy	377	YFLMSKTRNHS	387
Db	291	YFYTSDTIONS	301

RESULT 37

```

US-09-826-791-2
; Sequence 2, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-791-2

```

RESULT 38

```

US-09-828-478-5
; Sequence 5, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-5

```

RESULT 39

US-09-866-230-8
 ; Sequence 8, Application US/09866230
 ; Patent No. US20020150901A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Murphy, Andrew, et al.
 ; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Ma
 ; FILE REFERENCE: REG 771A
 ; CURRENT APPLICATION NUMBER: US/09/866,230
 ; CURRENT FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 60/207,725
 ; PRIOR FILING DATE: 2000-05-26
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 337
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-866-230-8

Query Match 17.3%; Score 372.5; DB 10; Length 337;
Best Local Similarity 30.2%; Pred. No. 8.5e-23;
Matches 103; Conservative 59; Mismatches 140; Indels 39; Gaps 12;

QY 92 ESASHLHVKNATMGYLTSSTLTKLPAIYLLVFGVGPANAVTLWMLF--FRTRSICTTV 149
DB 2 DETGNLTVSATCHDITDIDFRNOVYSTLYSMISVGVFGFGFYLYLKYTHKKS-AFQV 60

QY 150 FYTNLAIAIDFLFCVTLPFKIAYHLNGNNWVFGVLCRATTVIYFGNMYCSILLACISIN 209
DB 61 YMINLAVADLLCVCTPLRVVYVYHKGIMLFGDFCLRLSTYALYVNLVCSIFPMTAMSF 120

QY 210 RYLAIVHPFTYRGLPKHTALVTCGLWATVFLYMLPFFILKOEYVLVOPDITTCDDVHN 269
DB 121 RCIAIVFPVQNLNLTQKKARFVCGVGIWIEVILTSSPFLMAKPO--KDEKNNTKCFEPQ 178

QY 270 TCSSSPFQLYYFISLAFFGLIPFVLIYCYAAIIRTL-----NAYDHRWLWYVKAS 322
DB 179 DNGTKHVLVHVSU--FVGFIIPFVLIIVCYTMIILTLKSMKKNLSSHK-----KAI 232

QY 323 LLILVI---FTICFAPSNIILIIHHANYNNNT---DGLYFI---YLIALCLGSLNSCLD 373
DB 233 GMTWVTAAFLVSFMPHIORTI-HLHFLHNETKPCDSYLRMOKSVVITLSLAASNCPP 291

QY 374 PFLYFL-----MSKTRNH---STAYLTXXXNDLREQQ 403
DB 292 PLYFFSGGNFRKRLSTFRKHSLSSTVYVPRKASLPEKGE 332

RESULT 40

US-09-826-508-10
; Sequence 10, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; TITLE OF INVENTION: and Polynucleotides
; FILE REFERENCE: GP-70744USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-826-508-10

Query Match 17.1%; Score 368.5; DB 10; Length 358;
Best Local Similarity 29.3%; Pred. No. 1.9e-22;
Matches 89; Conservative 64; Mismatches 116; Indels 35; Gaps 9;

QY 100 KNATMGYLTSSTLTKLPAIYLLVFGVGPANAVTLWMLPFRTRSICTTVFY-TNLAIAD 158
DB 32 KNTT---LHNEFDTIPLVLYLIIFVASILLGLAVN-IPFHNRKTSFIFYLKNIWAD 87

QY 159 FLCFVTLPEKIAHLNGNNWVFGVLCRATTVIYFGNMYCSILLACISINRYLAIVHPF 218
DB 88 LIMTLTFPPFRIVHDAGFGPWYFKILCRYTSVLFYANMYTSIVFLGLISIDRYLKVKPKF 147

QY 219 TYRGLPKHTYALVTCGLWATVFLYMLPFFILKOEYVLVOPDITTCDDVHNTCESSSPFQ 278
DB 148 GDSRMYSITTKVLSCVWVIMAVLSLPNLTNG---QP---TEDNIHDCSKLSPLG 200

QY 279 LYFISLAFFG---FLIPFVLIYCYAAIIRTNAYDHRWL-----WYVKASLLILVI 328
DB 201 VKWHTAVTVNSCLFVAVLVILIGCYIAISRYIHKSSRQFISQSRKRKHQNSIRVVAV 260

QY 329 FTICFAPSNIILIIHHANYNNNTDG-----LYFIYLIALCLGSLNSCLDPFLYFL 379
DB 261 FFTCFULPHLCRI-----PFTFSLHDLRLDDESAQKILYYCKEITLFLSACNVCLDPIIYFF 316

QY 380 MSKT 383
DB 317 MCRS 320

Search completed: June 24, 2003, 12:18:37
Job time : 34.5399 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:01:21 ; Search time 18.0221 Seconds
(without alignments)
666.102 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIM.....AYLTXXNDLREQQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2151	99.7	408	2	US-08-742-440A-6
2	1433.5	66.9	407	2	US-08-742-440A-3
3	606.5	28.1	399	1	US-08-476-000-61
4	606.5	28.1	399	1	US-08-472-840-61
5	606.5	28.1	399	2	US-08-476-976-61
6	606.5	28.1	399	3	US-08-474-410-61
7	606.5	28.1	399	4	US-08-486-673B-61
8	605.5	28.1	395	1	US-08-097-938-2
9	605.5	28.1	395	1	US-08-476-000-2
10	605.5	28.1	395	1	US-08-472-840-2
11	605.5	28.1	395	2	US-08-476-976-2
12	605.5	28.1	395	3	US-08-474-410-2
13	603.5	28.0	395	1	US-08-097-938-5
14	603.5	28.0	395	1	US-08-476-000-5
15	603.5	28.0	395	1	US-08-472-840-5
16	603.5	28.0	395	2	US-08-476-976-5
17	603.5	28.0	395	3	US-08-474-410-5
18	603.5	28.0	395	4	US-08-486-673B-2
19	603.5	28.0	395	4	US-08-486-673B-5
20	585.5	27.1	398	1	US-08-097-938-4
21	585.5	27.1	398	1	US-08-476-000-4
22	585.5	27.1	398	1	US-08-472-840-4
23	585.5	27.1	398	2	US-08-476-976-4
24	585.5	27.1	398	3	US-08-474-410-4
25	585.5	27.1	398	4	US-08-486-673B-4
26	584.5	27.0	398	4	US-08-486-673B-6
27	583	27.0	394	2	US-08-742-440A-8

28	581.5	27.0	398	1	US-08-097-938-6	Sequence 6, Appli
29	581.5	27.0	398	1	US-08-476-000-6	Sequence 6, Appli
30	581.5	27.0	398	1	US-08-472-840-6	Sequence 6, Appli
31	581.5	27.0	398	2	US-08-476-976-6	Sequence 6, Appli
32	581.5	27.0	398	3	US-08-474-410-6	Sequence 6, Appli
33	577.5	26.8	397	4	US-08-486-673B-63	Sequence 63, Appli
34	569.5	26.4	397	1	US-08-476-000-63	Sequence 63, Appli
35	569.5	26.4	397	1	US-08-472-840-63	Sequence 63, Appli
36	569.5	26.4	397	2	US-08-476-976-63	Sequence 63, Appli
37	569.5	26.4	397	3	US-08-474-410-63	Sequence 63, Appli
38	526	24.4	425	1	US-08-097-938-7	Sequence 7, Appli
39	526	24.4	425	1	US-08-476-000-7	Sequence 7, Appli
40	526	24.4	425	1	US-08-472-840-7	Sequence 7, Appli
41	526	24.4	425	2	US-08-476-976-7	Sequence 7, Appli
42	526	24.4	425	3	US-08-474-410-7	Sequence 7, Appli
43	526	24.4	425	4	US-08-486-673B-7	Sequence 7, Appli
44	523	24.2	425	1	US-07-657-769B-69	Sequence 69, Appli
45	523	24.2	425	1	US-07-789-184-220	Sequence 220, App

ALIGNMENTS

RESULT 1
US-08-742-440A-6
; Sequence 6, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
US-08-742-440A-6

Query Match 99.7%; Score 2151; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Caps 0;


```
;
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-000-61

Query Match      28.1%; Score 606.5; DB 1; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPFALEGWT 79
Db 5 SLAWLLGGITLLAASVSCSRNTENLAPGRNNSKGRSLIGRLETOPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLSKLIPIAYLLVVFVGVVANATLWM 137
Db 52 GKGVPVEPGFSIDFSAS-----ILTGKLTTFVLPVVYIIIVFVIGLPSNGMALWI 101
QY 138 LPEFTSICITTFY-TNLAIADFLFCVTLPPFKIAYHLGNMNVFGEVLCRAITVIFVGNM 196
Db 102 FLFRTRKKHPAVIYMANLADLLSVIWFPLKISYHLGNMNVYGEALCKVLIGFFVGNM 161
QY 197 YCSILLACISINRYLAIIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
Db 162 YCSILFMTCLSVORYVIVNPM---GHPRKK-ANIAVGVSALWLLIFLVTIPLYVMKOT 217
QY 254 YVLVQPDITTCDDVHNTCESSPFQLY-----YFISLAPFGFLIPFVLIICYAAIIRT 307
Db 218 IYIPALNITTCDDV-----LPEEVLVGDVFNFLSLAIGVLFPAITASAYVLMIKT 270
QY 308 LNA-----YDHRWLWYVKASLLIIVIFTCFAPSNIILIIHHANYNNNTDGLYFIYLI 361
Db 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
Db 331 ALCLSTLNSCIDPFVYFVSKDFRDHA 357

RESULT 4
US-08-472-840-61
; Sequence 61, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
```

```
;
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-61

Query Match      28.1%; Score 606.5; DB 1; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPFALEGWT 79
Db 5 SLAWLLGGITLLAASVSCSRNTENLAPGRNNSKGRSLIGRLETOPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLSKLIPIAYLLVVFVGVVANATLWM 137
Db 52 GKGVPVEPGFSIDFSAS-----ILTGKLTTFVLPVVYIIIVFVIGLPSNGMALWI 101
QY 138 LPEFTSICITTFY-TNLAIADFLFCVTLPPFKIAYHLGNMNVFGEVLCRAITVIFVGNM 196
Db 102 FLFRTRKKHPAVIYMANLADLLSVIWFPLKISYHLGNMNVYGEALCKVLIGFFVGNM 161
QY 197 YCSILLACISINRYLAIIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
Db 162 YCSILFMTCLSVORYVIVNPM---GHPRKK-ANIAVGVSALWLLIFLVTIPLYVMKOT 217
QY 254 YVLVQPDITTCDDVHNTCESSPFQLY-----YFISLAPFGFLIPFVLIICYAAIIRT 307
Db 218 IYIPALNITTCDDV-----LPEEVLVGDVFNFLSLAIGVLFPAITASAYVLMIKT 270
QY 308 LNA-----YDHRWLWYVKASLLIIVIFTCFAPSNIILIIHHANYNNNTDGLYFIYLI 361
Db 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
Db 331 ALCLSTLNSCIDPFVYFVSKDFRDHA 357

RESULT 5
US-08-476-976-61
; Sequence 61, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
```

```
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-61

Query Match      28.1%; Score 606.5; DB 2; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPSALEGWT 79
DB 5 SLAWLLGGITLLAASVCSRTENLAPGRNNSKGRSLIGRLETPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLTKLIPAIYLLVVFVGVANAVTLWM 137
DB 52 GKGVPEVPGFSIDFSAS-----ILTGKLTTFVLPVVYIIVFVIGLPSNGMALWI 101
QY 138 LFRTSRISCTTVFY--TNLAIADELFCVTLPPKXAYHLGNWNVFGEVLCRAITVIFYGNM 196
DB 102 FLFRTRKXHPAVIYMANLADLLSVIWFPLKISYHLGNWNVYGEALCKVLIGFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
DB 162 YCSILFMTCLSVQRYWVWVNPW---GHPRKK-ANIAVGVSFLAIGVFLFPALLTASAYVLMIKT 217
QY 254 YVLVQPDITTCVDVNTCESSPFQLY-----YFISLAFPGFLIPFVLIICYAAIIT 307
DB 218 IYIPALNITTCVDV-----LPEEVLVGDVFNFLSLAIGVFLFPALLTASAYVLMIKT 270
QY 308 LNA-----YDHRWLVWVKASLLIIVFTICFAPSNIILIIHHANYVYNNNTDGLYFIYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNNLLLVVHYFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT--RNHS 387
DB 331 ALCLSTLNSCIDPFVYVYFVSKDFRDHA 357

RESULT 6
US-08-474-410-61
; Sequence 61, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-61

Query Match      28.1%; Score 606.5; DB 3; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPSALEGWT 79
DB 5 SLAWLLGGITLLAASVCSRTENLAPGRNNSKGRSLIGRLETPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLTKLIPAIYLLVVFVGVANAVTLWM 137
DB 52 GKGVPEVPGFSIDFSAS-----ILTGKLTTFVLPVVYIIVFVIGLPSNGMALWI 101
QY 138 LFRTSRISCTTVFY--TNLAIADELFCVTLPPKXAYHLGNWNVFGEVLCRAITVIFYGNM 196
DB 102 FLFRTRKXHPAVIYMANLADLLSVIWFPLKISYHLGNWNVYGEALCKVLIGFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
DB 162 YCSILFMTCLSVQRYWVWVNPW---GHPRKK-ANIAVGVSFLAIGVFLFPALLTASAYVLMIKT 217
QY 254 YVLVQPDITTCVDVNTCESSPFQLY-----YFISLAFPGFLIPFVLIICYAAIIT 307
DB 218 IYIPALNITTCVDV-----LPEEVLVGDVFNFLSLAIGVFLFPALLTASAYVLMIKT 270
QY 308 LNA-----YDHRWLVWVKASLLIIVFTICFAPSNIILIIHHANYVYNNNTDGLYFIYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNNLLLVVHYFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT--RNHS 387
DB 331 ALCLSTLNSCIDPFVYVYFVSKDFRDHA 357

RESULT 7
US-08-486-673B-61
; Sequence 61, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
```


REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-000-2

Query Match 28.1%; Score 605.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;

QY 30 LLLLLPTFCQSGMENDTNL--AKPTLPKIKTFRGAPPN---SPEEPFSALEGWGTGATIT 84
Db 17 VMILLRFLCTGRNNSKGRSLIGRLETOPPTITGKGVPEPGFSIDF----- 62
QY 85 VKIKPEESASHLVKNATMGYLTSSLSKLPALYLLVFPVGVDPANAVTLMLFRTS 144
Db 63 -----SAS-----ILTKLTTFVLEVVYIIIVFVGLPSNGMALWIFLRTTK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLPFKIAVHLGNWVFGVLCRATTVIFYGNMYCSILL 203
Db 105 KPAVIYMANLADLLSVIFWPKLSYHLGNWVYSEALCKVLGFFYGNMYCSILPM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VMAVFLYMLPFFILKQYIYLVQPD 260
Db 165 TCSLVQRYWVIVNPM---GHPKK-ANIAVGSLAIWLILFLVTIPLYVMKQIYIIPALN 220
QY 261 ITTCHDVNTCESSPPQLY-----YFISLAFPGFLIPPLVLIYCYAIIITLNA--- 310
Db 221 ITTCHDV-----LPEEVLVGMFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMD 273
QY 311 --YDHRWLWVVKASLLILVIFTCFAPSNIILIIHHANYNNYNTDGLYFYLIYALCLGSL 368
Db 274 ESENKQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTORQSHVYALVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
Db 334 NSCIDPFVYFVSKDRDHA 353

RESULT 10
US-08-472-840-2
Sequence 2, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/472,840
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-2

Query Match 28.1%; Score 605.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;

QY 30 LLLLLPTFCQSGMENDTNL--AKPTLPKIKTFRGAPPN---SPEEPFSALEGWGTGATIT 84
Db 17 VMILLRFLCTGRNNSKGRSLIGRLETOPPTITGKGVPEPGFSIDF----- 62
QY 85 VKIKPEESASHLVKNATMGYLTSSLSKLPALYLLVFPVGVDPANAVTLMLFRTS 144
Db 63 -----SAS-----ILTKLTTFVLEVVYIIIVFVGLPSNGMALWIFLRTTK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLPFKIAVHLGNWVFGVLCRATTVIFYGNMYCSILL 203
Db 105 KPAVIYMANLADLLSVIFWPKLSYHLGNWVYSEALCKVLGFFYGNMYCSILPM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VMAVFLYMLPFFILKQYIYLVQPD 260
Db 165 TCSLVQRYWVIVNPM---GHPKK-ANIAVGSLAIWLILFLVTIPLYVMKQIYIIPALN 220
QY 261 ITTCHDVNTCESSPPQLY-----YFISLAFPGFLIPPLVLIYCYAIIITLNA--- 310
Db 221 ITTCHDV-----LPEEVLVGMFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMD 273
QY 311 --YDHRWLWVVKASLLILVIFTCFAPSNIILIIHHANYNNYNTDGLYFYLIYALCLGSL 368
Db 274 ESENKQRAIRLIITVLAMYFICFAPSNLLLVVHYFLIKTORQSHVYALVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
Db 334 NSCIDPFVYFVSKDRDHA 353

RESULT 11
US-08-476-976-2
Sequence 2, Application US/08476976
Patent No. 5874400
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/476,976
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/390,301
;; FILING DATE: 25-JAN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ADLER, REID G.
;; REGISTRATION NUMBER: 30,988
;; REFERENCE/DOCKET NUMBER: 2803-0006.20
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 887-1500
;; TELEFAX: (202) 887-0763
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 395 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-476-976-2

Query Match 28.1%; Score 605.5; DB 2; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;
QY 30 LLLLLPTFCQSGMENDTNL--AKPTLPKTRGAPN---SPEEPFSALEGTGATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLIGRLETPPTITGKGVPEPGFSIDEP----- 62
QY 85 VKIKPEESASHLHVKNATMGYLTSLSLTKLIPAIYLLVFWGVGPANAVTLMLFFRTS 144
DB 63 -----SAS-----ILTGKLTTFVLPVYIIVFVIGLPSNGMALWIFLRTKK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLFPKIAHYLHNGNWNVFGVLCRAITVIFYGNYCISILL 203
DB 105 KHPAVIYMANLADLLSVIFWFLPKISYHLHNGNWNVYGEALCKVLIGFFYGNMYCSILFM 164
QY 204 ACISINRYLAIVHPFTYRGKPKHYALVTGCL--VMATVFLYMLPFILKQEVYLVQPD 260
DB 165 TCSLVQRYWVIVNPM---GHPKK-ANIAVGSLAIWLLIFLVTIPLYVWKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSPPOLY-----YFISLAFGFLIPFVLIYCYAAIIRTNA---- 310
DB 221 ITTCHDV-----LPEEVLGDMFNFLSLAIGVFLPALLTASAYVLMKILRSSAMD 273
QY 311 --YDHRWLWYVKASLLILVIFTCFAPSNIILIIHANYNNYNTDGLYFIYLIACLGSL 368
DB 274 EHSNKRQRAIRLIITVLAIFYCFAPSNNLLVHVFLIKTORQSHVYALYVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYVYVSKDFRHA 353

RESULT 12
US-08-474-410-2
; Sequence 2, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

;; ZIP: 20006-1812
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,410
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/390,301
;; FILING DATE: 25-JAN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ADLER, REID G.
;; REGISTRATION NUMBER: 30,988
;; REFERENCE/DOCKET NUMBER: 2803-0006.20
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 887-1500
;; TELEFAX: (202) 887-0763
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 395 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-474-410-2

Query Match 28.1%; Score 605.5; DB 3; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;
QY 30 LLLLLPTFCQSGMENDTNL--AKPTLPKTRGAPN---SPEEPFSALEGTGATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLIGRLETPPTITGKGVPEPGFSIDEP----- 62
QY 85 VKIKPEESASHLHVKNATMGYLTSLSLTKLIPAIYLLVFWGVGPANAVTLMLFFRTS 144
DB 63 -----SAS-----ILTGKLTTFVLPVYIIVFVIGLPSNGMALWIFLRTKK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLFPKIAHYLHNGNWNVFGVLCRAITVIFYGNYCISILL 203
DB 105 KHPAVIYMANLADLLSVIFWFLPKISYHLHNGNWNVYGEALCKVLIGFFYGNMYCSILFM 164
QY 204 ACISINRYLAIVHPFTYRGKPKHYALVTGCL--VMATVFLYMLPFILKQEVYLVQPD 260
DB 165 TCSLVQRYWVIVNPM---GHPKK-ANIAVGSLAIWLLIFLVTIPLYVWKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSPPOLY-----YFISLAFGFLIPFVLIYCYAAIIRTNA---- 310
DB 221 ITTCHDV-----LPEEVLGDMFNFLSLAIGVFLPALLTASAYVLMKILRSSAMD 273
QY 311 --YDHRWLWYVKASLLILVIFTCFAPSNIILIIHANYNNYNTDGLYFIYLIACLGSL 368
DB 274 EHSNKRQRAIRLIITVLAIFYCFAPSNNLLVHVFLIKTORQSHVYALYVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYVYVSKDFRHA 353

RESULT 13
US-08-097-938-5
; Sequence 5, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,938
FILING DATE: 26-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22803-20006.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-097-938-5

Query Match 28.0%; Score 603.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;

QY 30 LLLLPFTCSGMENTNNL--AKPTLPKTRGAPPN---SPEEPFSALEGTGATIT 84
DB 17 VMILLRPLCTGRNNSKGRSLIGRLTQPTITGKVPVEPGFSIDF----- 62

QY 85 VKIKPEESASHLHVKNATGYLTSSLTSLKLIPIAYLLVGVGPANAVTLWMLFRTS 144
DB 63 -----SAS-----ILTGKLTTFVPPVYIIIVGLPSNGMALIFLRTTK 104

QY 145 ICTTVFY-TNLAIADFLFCVTLFPFKIAYHLNGNNWVFGVLCRATTVIFYGNNMYSILL 203
DB 105 KHPAVIYMANLADLLSVLPFKISYHLGNNWVGEALCKVLGFFYGNMYSILFM 164

QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTCL---VWATVFLMPLFFILKQBYLVQPD 260
DB 165 TCLSVQRYWVIVNPM---GHPRKK-ANIAVGSLSLAIWLLIFLVTIPLYNMKOTIYIPALN 220

QY 261 ITTCHDVNTCESSSPFOLY-----YFISLAFPGFLIPFVLIICYAAIIRTNA---- 310
DB 221 ITTCHDV-----LPEEVLVGMFNFLSLAIGVLPALLTASAYLVMIKTLRSSAMD 273

QY 311 --YDHRWLWYKASLLILVIFTCFAPSNIILIIHHANNYYNNNDGLYFIYIALCLGSL 368
DB 274 EHSEKKQRAIRLIITVLAMYFCFAPSNLLLVVHYFLIKTORQSHVYALYVALCLSTL 333

QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYFVSKDFRHA 353

RESULT 14
US-08-476-000-5
Sequence 5, Application US/08476000
Patent No. 5716789
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
TITLE OF INVENTION: 63
NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-476-000-5

Query Match 28.0%; Score 603.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;

QY 30 LLLLPFTCSGMENTNNL--AKPTLPKTRGAPPN---SPEEPFSALEGTGATIT 84
DB 17 VMILLRPLCTGRNNSKGRSLIGRLTQPTITGKVPVEPGFSIDF----- 62

QY 85 VKIKPEESASHLHVKNATGYLTSSLTSLKLIPIAYLLVGVGPANAVTLWMLFRTS 144
DB 63 -----SAS-----ILTGKLTTFVPPVYIIIVGLPSNGMALIFLRTTK 104

QY 145 ICTTVFY-TNLAIADFLFCVTLFPFKIAYHLNGNNWVFGVLCRATTVIFYGNNMYSILL 203
DB 105 KHPAVIYMANLADLLSVLPFKISYHLGNNWVGEALCKVLGFFYGNMYSILFM 164

QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTCL---VWATVFLMPLFFILKQBYLVQPD 260
DB 165 TCLSVQRYWVIVNPM---GHPRKK-ANIAVGSLSLAIWLLIFLVTIPLYNMKOTIYIPALN 220

QY 261 ITTCHDVNTCESSSPFOLY-----YFISLAFPGFLIPFVLIICYAAIIRTNA---- 310
DB 221 ITTCHDV-----LPEEVLVGMFNFLSLAIGVLPALLTASAYLVMIKTLRSSAMD 273

QY 311 --YDHRWLWYKASLLILVIFTCFAPSNIILIIHHANNYYNNNDGLYFIYIALCLGSL 368
DB 274 EHSEKKQRAIRLIITVLAMYFCFAPSNLLLVVHYFLIKTORQSHVYALYVALCLSTL 333

QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYFVSKDFRHA 353

RESULT 15
US-08-472-840-5
Sequence 5, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,840
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US/08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-472-840-5

Query Match 28.0%; Score 603.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;

QY 30 LLLLPFTFCOSGMENDTNL--AKPTLPKTPRGAPPN---SPEEPFSALEGWGTATIT 84
Db 17 VMILLRFLCTGRNNSKGRSLIGRLETPPTITGKGVPEPGFSIDEP-----62
QY 85 VKIKPEESASHLVKNATMGYLTSSLSKLIPIAYLLVVFVGVGPANAVTLMWLPFRTRS 144
Db 63 -----SAS-----ILTGKLTTFVLPVWYIIVFVIGLPSNGMALWFLPRTKK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLFPKTAHYHLNGNNWVGEVLCRAITTVIFGNNMYSILL 203
Db 105 KHPAVIYMANLADALLSVIFWFLPKISYHLHGNWVYGEALCKVLGFFYGNMYSILFM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTGCL---VMATVFLMPLPFILKQBYLVQPD 260
Db 165 TCSLVQRWYVWVNM---GHPKK-ANIAVGVSALWLLIFLVTIPLYVWKOTIYIPALN 220
QY 261 ITTCHDVHNTCESSPPQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA---- 310
Db 221 ITTCHDV-----LPEEVLVGMENYFSLAIGVFLPALLTASAYVLMIKTLRSSAMD 273
QY 311 --YDHRWLWTVKASLLILVITFCFAPSNIILIHANYYNNYNTDGLYFIYLALCLGSL 368
Db 274 EHSEKKRQRAIRLIITVLAIFYCFAPSNLLLVVHYFLIKTORQSHVYALVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
Db 334 NSCIDPFVYVFSKDFRDA 353

RESULT 16
US-08-476-976-5
Sequence 5, Application US/08476976
Patent No. 5874400
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,976
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-476-976-5

Query Match 28.0%; Score 603.5; DB 2; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;

QY 30 LLLLPFTFCOSGMENDTNL--AKPTLPKTPRGAPPN---SPEEPFSALEGWGTATIT 84
Db 17 VMILLRFLCTGRNNSKGRSLIGRLETPPTITGKGVPEPGFSIDEP-----62
QY 85 VKIKPEESASHLVKNATMGYLTSSLSKLIPIAYLLVVFVGVGPANAVTLMWLPFRTRS 144
Db 63 -----SAS-----ILTGKLTTFVLPVWYIIVFVIGLPSNGMALWFLPRTKK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLFPKTAHYHLNGNNWVGEVLCRAITTVIFGNNMYSILL 203
Db 105 KHPAVIYMANLADALLSVIFWFLPKISYHLHGNWVYGEALCKVLGFFYGNMYSILFM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTGCL---VMATVFLMPLPFILKQBYLVQPD 260
Db 165 TCSLVQRWYVWVNM---GHPKK-ANIAVGVSALWLLIFLVTIPLYVWKOTIYIPALN 220
QY 261 ITTCHDVHNTCESSPPQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA---- 310
Db 221 ITTCHDV-----LPEEVLVGMENYFSLAIGVFLPALLTASAYVLMIKTLRSSAMD 273
QY 311 --YDHRWLWTVKASLLILVITFCFAPSNIILIHANYYNNYNTDGLYFIYLALCLGSL 368
Db 274 EHSEKKRQRAIRLIITVLAIFYCFAPSNLLLVVHYFLIKTORQSHVYALVALCLSTL 333

;; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
;; FILE REFERENCE: 44481-5006-08-US
;; CURRENT APPLICATION NUMBER: US/08/486, 673B
;; PRIOR FILING DATE: 1995-06-07
;; PRIOR APPLICATION NUMBER: US 08/097,938
;; PRIOR FILING DATE: 1993-07-26
;; PRIOR APPLICATION NUMBER: PCT/US94/08536
;; PRIOR FILING DATE: 1994-07-26
;; NUMBER OF SEQ ID NOS: 63
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 5
;; LENGTH: 395
;; TYPE: PRT
;; ORGANISM: Mus musculus
;; FEATURE:
;; OTHER INFORMATION: Protein sequence of C140 receptor
US-08-486-673B-5

Query Match 28.0%; Score 603.5; DB 4; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;
QY 30 LLLLLLTCQSGMENDTNL--AKPTLPKIFRGGAPPN---SPEEPFSALEGWTGATIT 84
Db 17 VMILLRFLCTGRNRSKRSGLRLETPPTTGKGVPEPGFSIDEP----- 62
QY 85 VKIKCPESASHLHVKNATWGYLTSSLSLTKLIPAIYLLVFVGVGPANAVTLMLFRTS 144
Db 63 -----SAS-----ILTGKLTTFVLPVVIIVFVIGLPSNGMALWFLRFTTK 104
QY 145 ICTTVFY-TNLATADFLFCVTLFPKIAHYHLNGNNWVGEVLCRATTVIFYGNMYC 203
Db 105 KHPAVIYMANLADLLSVIWFPLKIAHYHGNWIIYGEALCNVLIGFFYGNMYC 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKQEVYLVQPD 260
Db 165 TCSVORYWVIVNPM---GHPKK-ANIAVGSLSLAILLFLVTLPIYVKNQTIYIPALN 220
QY 261 ITTCHDVNTCESSPPQLY-----YFISLAFGLFPLIPVLIYCYAAIIRTNA--- 310
Db 221 ITTCHDV-----LPEEVLVGMENYFSLAIGVFLFPALLTASAYVLMIKLRSAMD 273
QY 311 --YDHRWLWTVKASLLILVITFCFAPSNIILIIHHANYNNYNTDGLYFIYALCIGSL 368
Db 274 EHSEKQRQRAIRLIITVLAHYFCFAPSNIILVHVHFLIKTORQSHVYALYVALCLSTL 333
QY 369 NSCLDPLFLYMSKT-RNHS 387
Db 334 NSCIDPFPVYFVSKDFRHA 353

RESULT 20
US-08-097-938-4
; Sequence 4, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995

;; APPLICATION NUMBER: US/08/097,938
;; FILING DATE: 26-JUL-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MURASHIGE, KATE H.
;; REGISTRATION NUMBER: 29,959
;; REFERENCES/DOCKET NUMBER: 22803-20006.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 887-1500
;; TELEFAX: (202) 887-0763
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 398 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-097-938-4
Query Match 27.1%; Score 585.5; DB 1; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;
QY 80 GATITVVKIKCPESASHLHVKNATWGYLTSSLSLTKLIPAIYLLVFVGVGPANAVTLMLF 139
Db 53 GVTETVFSVDIFSAS-----VLTKLTTVLPPIVYTIIVFVGLPSNGMALWVFL 102
QY 140 FRTSRICITTVFY-TNLATADFLFCVTLFPKIAHYHLNGNNWVGEVLCRATTVIFYGNMYC 198
Db 103 FRTKXHPAVIYMANLADLLSVIWFPLKIAHYHGNWIIYGEALCNVLIGFFYGNMYC 162
QY 199 SILLIACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEVYLVQ 258
Db 163 SILFMTCLSVQRYWVIVNPMGH-SRKKANIAIGSLAILLWLLVTLPIYVVKQTIPIA 221
QY 259 PDIITCHDVNTCESSPPQLY-----YFISLAFGLFPLIPVLIYCYAAIIRTNA--- 310
Db 222 LNTITCHDV-----LPEQLLVGMENYFSLAIGVFLFPALLTASAYVLMIRLSSA 274
QY 311 ---YDHRWLWTVKASLLILVITFCFAPSNIILIIHHANYNNYNTDGLYFIYAL 363
Db 275 MDENSEKKRKAIRKILVTVLAMYLCFTPSNLLLVH---YFLIKSQGSHVYALVIVAL 331
QY 364 CIGLSNCLDPLFLYMSKT-RNHS-TAYLTXXNDLRE 400
Db 332 CLSTLNSCIDPFPVYFVSHDFRDAKNAKLLCRSVRTKQ 370

RESULT 21
US-08-476-000-4
; Sequence 4, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-000-4

Query Match      27.1%; Score 585.5; DB 1; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY      80 GATITVKIKCPESASHLHVNKATMGVLTSSLSKLIPIALYLLVFVGVGPANAVTLWMLF 139
DB      53 GVTVEVFSVDFESAS-----VLTKGLTTFVLPVYTVFVVGLPNGMALWVFL 102

QY      140 FRTSICITVFY-TNLAIDFLFCVTLFPKIAVHLNGNNWVGEVLCRATTVIFYGNMYC 198
DB      103 FRTKKHPAIVYMANLADLLSVIWPFLKIAYHIGNNNIYGEALCNVLIGFFYGNMYC 162

QY      199 SILLACISINRYLAIVHPFTYRGPKHTYALVTCGLVWATVFLYMLPFFILKOEYVLVQ 258
DB      163 SILFMTCLSVQRVYVWVPMGH-SRKKANAIAGISLAILWLLLTPIYVVKQTIFIPA 221

QY      259 PDITTCCHDVNTCESSPPFLY-----YFISLAFGFLIPFLVLIYCYAAIIRTLNA-- 310
DB      222 LNIITCHDV-----LPEQLVGDVFNFLSALGVLPALFASAYVLMIRLSSA 274

QY      311 ----YDHRWLWYVKASLLILVITFCAPSNIILIIHHANYNNNDG---LYFIYLIAL 363
DB      275 MDENSEKKRAIKLIIVTLVAMYLICFTPSNLLLVH---YFLIKSQGSHVVALYVAL 331

QY      364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
DB      332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 22
US-08-472-840-4
; Sequence 4, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
```

```
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-4

Query Match      27.1%; Score 585.5; DB 1; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY      80 GATITVKIKCPESASHLHVNKATMGVLTSSLSKLIPIALYLLVFVGVGPANAVTLWMLF 139
DB      53 GVTVEVFSVDFESAS-----VLTKGLTTFVLPVYTVFVVGLPNGMALWVFL 102

QY      140 FRTSICITVFY-TNLAIDFLFCVTLFPKIAVHLNGNNWVGEVLCRATTVIFYGNMYC 198
DB      103 FRTKKHPAIVYMANLADLLSVIWPFLKIAYHIGNNNIYGEALCNVLIGFFYGNMYC 162

QY      199 SILLACISINRYLAIVHPFTYRGPKHTYALVTCGLVWATVFLYMLPFFILKOEYVLVQ 258
DB      163 SILFMTCLSVQRVYVWVPMGH-SRKKANAIAGISLAILWLLLTPIYVVKQTIFIPA 221

QY      259 PDITTCCHDVNTCESSPPFLY-----YFISLAFGFLIPFLVLIYCYAAIIRTLNA-- 310
DB      222 LNIITCHDV-----LPEQLVGDVFNFLSALGVLPALFASAYVLMIRLSSA 274

QY      311 ----YDHRWLWYVKASLLILVITFCAPSNIILIIHHANYNNNDG---LYFIYLIAL 363
DB      275 MDENSEKKRAIKLIIVTLVAMYLICFTPSNLLLVH---YFLIKSQGSHVVALYVAL 331

QY      364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
DB      332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 23
US-08-476-976-4
; Sequence 4, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```


; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-4

Query Match 27.1%; Score 585.5; DB 2; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGYLTSSLTSLKIPAIYLLVVFVGVGPANAVTLWMLF 139
DB 53 GVTVEVFSVDFESAS-----VLTGKLTTFVLPVIVTVFVGLPSNGMALWVFL 102

QY 140 FRTRISCTTVFY-TNLAIADFLFCVTLPPFKIAHYLNGNNWVGEVLCRATTVIFYGNMYC 198
DB 103 FRTKKHPAIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYC 162

QY 199 SIILLACISINRYLAIVHPFTYRGPKHYALVTCGLVWATVFLYMLPFFILKQEYVLVQ 258
DB 163 SILFMTCLSVQRVYVWPNMGH-SRKANIAIGISLAIWLLILVTLPIYVVKQTIFIPA 221

QY 259 PDITTCCHDVNTCESSPPFOLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA-- 310
DB 222 LNIITTCADV-----LPEQLLVGDMFNFLSLAIGVFLPAPLTASAVLMIRLSSA 274

QY 311 ----YDHRWLWYKASLLILVITFCAPSNIILIIHANYNNYNTDG----LYFIYLIAL 363
DB 275 MDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGQSHVVALYVAL 331

QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 24
US-08-474-410-4
; Sequence 4, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-4

Query Match 27.1%; Score 585.5; DB 3; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGYLTSSLTSLKIPAIYLLVVFVGVGPANAVTLWMLF 139
DB 53 GVTVEVFSVDFESAS-----VLTGKLTTFVLPVIVTVFVGLPSNGMALWVFL 102

QY 140 FRTRISCTTVFY-TNLAIADFLFCVTLPPFKIAHYLNGNNWVGEVLCRATTVIFYGNMYC 198
DB 103 FRTKKHPAIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYC 162

QY 199 SIILLACISINRYLAIVHPFTYRGPKHYALVTCGLVWATVFLYMLPFFILKQEYVLVQ 258
DB 163 SILFMTCLSVQRVYVWPNMGH-SRKANIAIGISLAIWLLILVTLPIYVVKQTIFIPA 221

QY 259 PDITTCCHDVNTCESSPPFOLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA-- 310
DB 222 LNIITTCADV-----LPEQLLVGDMFNFLSLAIGVFLPAPLTASAVLMIRLSSA 274

QY 311 ----YDHRWLWYKASLLILVITFCAPSNIILIIHANYNNYNTDG----LYFIYLIAL 363
DB 275 MDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGQSHVVALYVAL 331

QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 25
US-08-486-673B-4
; Sequence 4, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens

```
US-08-486-673B-4
Query Match      27.1%; Score 585.5; DB 4; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGVLTSSLTKLIPALYLLVFFVGVGPANAVTLWMLF 139
DB 53 GVIVEIVFSDERSAS-----VLTGKLTIVFLPIVYIIVFVGVLPNGMALVFL 102
QY 140 FRTSICITVYF-TNLAIADFLFCVTLPPFKIAHNLGNWVFGVLCRATTVIFYGNMYC 198
DB 103 FRTKKKHPAIVYMANIALADLLSVIWFPLKIAYHIHGNWNIYGEALCNVLIGFFYGNMYC 162
QY 199 SILLACISINRYLAIVHPYTYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVYLVQ 258
DB 163 SILFMTCLSVQRVYVWVNPNGH-SRKKANIAIGISLAIMLLILLVTLPIVYVQOTIFIPA 221
QY 259 PDITTCCHDVNTCESSPPFQLY-----YFISLAFFGFLIPFVLIIVCYAAIIRTLNA-- 310
DB 222 LNTTCHDV-----LPEQLLVGDMENYFLSLAIGVFLPAFLTASAYVLMIRLSSA 274
QY 311 -----YDHRWLWYVKASLLILVITFCPAPSNIIILHHANYNNYNDG---LYFIYLIAL 363
DB 275 MDENSEKKRAIKLIIVTLAMYLICFTPSNLLLVH---YFLIKSQGQSHVYALYIVAL 331
QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 27
US-08-742-440A-8
; Sequence 8, Application US/08742440A
; Patent No. 582014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-440A-8

Query Match      27.0%; Score 583; DB 2; Length 394;
Best Local Similarity 35.1%; Pred. No. 5.6e-40;
Matches 136; Conservative 65; Mismatches 139; Indels 48; Gaps 13;

QY 28 AGLLLLLPTFCGSGMENDTNLAKPLTKTRGAPPNSFEFFPSALSGWGTGATITVKI 87
DB 12 AAIIAASLSC-SGTIGTNRSSKGRSLIGKVDGT-----SHVTG-KGVTVETVF 59
QY 88 KCPEBSASHLVKNATMGVLTSSLTKLIPALYLLVFFVGVGPANAVTLWMLPRTSICT 147
DB 60 SVDEFSA-----VLTGKLTIVFLPIVYIIVFVGVLPNGMALVFLRTKKHP 109
QY 148 TVFY-TNLAIADFLFCVTLPPFKIAHNLGNWVFGVLCRATTVIFYGNMYCSILLACI 206
```

Db 110 AVTYMANLADLLSVIWPFLKTIAYHGHNNWYIGALCNVLIGFFYGNMYCSILEWTC 169
 Qy 207 SINRYLAIPHFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQPDITTC 266
 Db 170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIMWLLLVTPLYVVKQTIPIPALNITTC 228
 Qy 267 VHTCSESSPQLYY---FISLAFPGFLIPFVLIYCYAAIIRTLNA-----YDHRWLW 317
 Db 229 V-----LPEQLLVGDPFLSLAIGVFLPFAFLTASAYVLMIRLSSAMDENSEKKRR 281
 Qy 318 YVKASLLILVITFCAPSNIILIIHHANYNNYNTDG---LYFIYLIALCLGSLNSCLDP 374
 Db 282 AILKIVTLVAMYLICFTPSNLLLVH---YFLKSGQSHVAYIYVALCLSTLNSCIDP 338
 Qy 375 FLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
 Db 339 FVYFVSHDFRDHAKNALLCRSVRTVKQ 366

RESULT 28

US-08-097-938-6
 ; Sequence 6, Application US/08097938
 ; Patent No. 5629174
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
 ; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/097,938
 ; FILING DATE: 26-JUL-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22803-20006.00
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-097-938-6

Query Match 27.0%; Score 581.5; DB 1; Length 398;
 Best Local Similarity 36.0%; Pred. No. 7.5e-40;
 Matches 122; Conservative 65; Mismatches 113; Indels 39; Gaps 10;

Qy 80 GATITVKIKCPESASHLVKNATMGVLTSSLTSLKIPAIYLLVVFVGVGPANAVTLWMLF 139
 Db 53 GVIVEIVFSVDFESAS-----VLTKGLTTVFLPIVYIIVFVVGLPNSGMALWFL 102
 Qy 140 FRTSICITTVFY-TNLAIADFLFCVTLPPKIAHYLNGNNWVFGVLCRATTVIFYGNMYC 198
 Db 103 FRTKKHPAVIYMANLADLLSVIWPFLKTIAYHGHNNWYIGALCNVLIGFFYGNMYC 162
 Qy 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQ 258

Db 163 SILFMTCLSVQRYWVIVNPMGH-SRKKANIAIGISLAIMWLLLVTPLYVVKQTIPIPA 221
 Qy 259 PDITTCADVHTNCSESSPQLYY---YFISLAFPGFLIPFVLIYCYAAIIRTLNA-- 310
 Db 222 LMITTCADV-----LPEQVLVGMFNFYLSLAIGVFLPFAFLTASAYVLMIRLSSA 274
 Qy 311 ----YDHRWLWYVKASLLILVITFCAPSNIILIIHHANYNNYNTDG---LYFIYLIAL 363
 Db 275 MDENSEKKRKAUKLIVTVLAMYLCIFSPNLLLVH---YFLKSGQSHVAYIYVAL 331
 Qy 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
 Db 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 29

US-08-476-000-6
 ; Sequence 6, Application US/08476000
 ; Patent No. 5716789
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,000
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-476-000-6

Query Match 27.0%; Score 581.5; DB 1; Length 398;
 Best Local Similarity 36.0%; Pred. No. 7.5e-40;
 Matches 122; Conservative 65; Mismatches 113; Indels 39; Gaps 10;

Qy 80 GATITVKIKCPESASHLVKNATMGVLTSSLTSLKIPAIYLLVVFVGVGPANAVTLWMLF 139
 Db 53 GVIVEIVFSVDFESAS-----VLTKGLTTVFLPIVYIIVFVVGLPNSGMALWFL 102
 Qy 140 FRTSICITTVFY-TNLAIADFLFCVTLPPKIAHYLNGNNWVFGVLCRATTVIFYGNMYC 198
 Db 103 FRTKKHPAVIYMANLADLLSVIWPFLKTIAYHGHNNWYIGALCNVLIGFFYGNMYC 162

Db 103 FRTKKHPAVIYMANLADLLSVIWFPLKIAIYHGNHNNIYGEALCNVLIGFFYGNMYC 162
Qy 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQ 258
Db 163 SILFMTCLSVORYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVITIPLYVVKQTIFIPA 221
Qy 259 PDITTCVHNTCESSPPOLY-----YFISLAFFGLIPFVLIYCYAAIIRTNA-- 310
Db 222 LNTITTCVH-----LPEQVLVGDVGFVFLSIAIGVFLPAPFLTASAYVLMIRLSSA 274
Qy 311 ----YDHRMLWYVKASLLIIVITFCFAPSNIILIIHANYNNNTDG---LYFIYLI 363
Db 275 MDENSEKKRRAIKLIVTVLAMYLCIFIPSNLLVH---YFLIKSQGQSHVYALIVAL 331
Qy 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTCKXNDLRE 400
Db 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 32
US-08-474-410-6
; Sequence 6, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1912
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/474,410
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-474-410-6

Query Match 27.08; Score 581.5; DB 3; Length 398;
Best Local Similarity 36.08; Pred. No. 7.5e-40;
Matches 122; Conservative 65; Mismatches 113; Indels 39; Gaps 10;

Qy 80 GATITVIKCPESASHLVKNATMGVLTSSLTKLIPALYLVFVGVGPANAVTLWMLF 139
Db 53 GVIVEIVFSDFSAS-----VLTGKLTIVFLPIYIVFVVGVLPSNGMALVFL 102
Qy 140 FRTSICITVY-TNLAIDFLFCVTLPPFKIAIYHNLGNHNNVFGVLCRATTIVIFYGNMYC 198

Db 103 FRTKKHPAVIYMANLADLLSVIWFPLKIAIYHGNHNNIYGEALCNVLIGFFYGNMYC 162
Qy 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQ 258
Db 163 SILFMTCLSVORYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVITIPLYVVKQTIFIPA 221
Qy 259 PDITTCVHNTCESSPPOLY-----YFISLAFFGLIPFVLIYCYAAIIRTNA-- 310
Db 222 LNTITTCVH-----LPEQVLVGDVGFVFLSIAIGVFLPAPFLTASAYVLMIRLSSA 274
Qy 311 ----YDHRMLWYVKASLLIIVITFCFAPSNIILIIHANYNNNTDG---LYFIYLI 363
Db 275 MDENSEKKRRAIKLIVTVLAMYLCIFIPSNLLVH---YFLIKSQGQSHVYALIVAL 331
Qy 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTCKXNDLRE 400
Db 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 33
US-08-486-673B-63
; Sequence 63, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-486-673B-63

Query Match 26.88; Score 577.5; DB 4; Length 397;
Best Local Similarity 34.4%; Pred. No. 1.6e-39;
Matches 136; Conservative 64; Mismatches 144; Indels 51; Gaps 13;

Qy 28 AGLLLLPTFCOSGMENDTNLAKPTLPKTRGAPPNSFEFPFSALESGWTGATITVKI 87
Db 12 AAILLAASLSC-SGTIQGTNRSSKRSGLIGKVDGT-----SHVTG-KGVTVTVPF 59
Qy 88 KCPEESASHLVKNATMGVLTSSLTKLIPALYLVFVGVGPANAVTLWMLPRTSIC 147
Db 60 SVDEFSAS-----VLAKLTIVFLPIYIVFVVGVLPSNGMALVFLFRTKKHP 109
Qy 148 TVFY-TNLAIDFLFCVTLPPFKIAIYHNLGNHNNVFGVLCRATTIVIFYGNMYCILLACI 206
Db 110 AVIYMANLADLLSVIWFPLKIAIYHGNHNNIYGEALCNVLIGFFYGNMYCSILPMTCL 169
Qy 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQDDITTC 266
Db 170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVITIPLYVVKQTIFIPALNITTC 228
Qy 267 VHTCESSPPOLY-----YFISLAFFGLIPFVLIYCYAAIIRTNA-----YDHR 314
Db 229 V-----LPEQLLVGDVGFVFLSIAIGVFLPAPFLTASAYVLMIRLSSANDENSEKK 281
Qy 315 WLWYVKASLLIIVITFCFAPSNIILIIHANYNNNTDG---LYFIYLIACLSLNSC 371
Db 282 RRAIKLIVTVLGMVLCIFTPSNLLVH---YFLIKSQGQSHVYALIVALCLSLNSC 338
Qy 372 LDPFLYFLMSKT-RNHS-TAYLTCKXNDLREOQOP 404

Db 339 IDPFVYFVSHDRHAKNALLCRSVRTVKQMVP 373

RESULT 34

US-08-476-000-63
; Sequence 63, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-000-63

Query Match 26.4%; Score 569.5; DB 1; Length 397;
Best Local Similarity 34.2%; Pred. No. 7.1e-39;
Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

Qy 28 AGLLLLPTFCQSGMENDTNNAKPTLPKTFRGAPPNSFEFFPSALEGWTGATITVKI 87
Db 12 AAIIAASLSC-SGTIOQTNRSKGRSLGKVDGT-----SHVTG-KGVTVTVF 59
Qy 88 KCPEESASHLVKNATMGVLTSSLTSLIPAIYLVFVGVGPANAVTLWMLFRTSICT 147
Db 60 SVDEFSAS-----VLAKLTIVFLPIVYIVFAVGLPSNGMALVFLFRTKKHP 109
Qy 148 TVFY-TNLAIADFLCVTLFPFKIAYHLNGNNWVGEVLCRATTIVFYGNMYCSILLACI 206
Db 110 AVIYMANLADLLSVIWPFLPKIAYHIGNNWIYGEALCNVLGFFYRNMYCSILFMTCL 169
Qy 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQYLVQPDITTC 266
Db 170 SVORYWIVNPMGH-SRKKANIAIGISLAIWLLTLVTLPIYVVKOTIFIPALNITTC 228
Qy 267 VHTNCSSSPOLY-----YFISLAPFFGLPVLIIYCYAAILRTLNA-----YDHR 314
Db 229 V-----LPEQLLVGMFMFYSLAIGVFLFAFLTASAYVLMIKRLSRSDENSEKK 281
Qy 315 WLWVYKASLLILVIFTICFAPSNIILIIHHYNNYNTDG---LYFIYIALCLGSLNSC 371

Db 282 RKRAIKLIVTLGMYLICFTPSNLLLVH---YFLIKSQGSHVYALVALCLSTLNSC 338
Qy 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREQQP 404
Db 339 IDPFVYFVSHDRHAKNALLCRSVRTVKQMVP 373

RESULT 35

US-08-472-840-63
; Sequence 63, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-63

Query Match 26.4%; Score 569.5; DB 1; Length 397;
Best Local Similarity 34.2%; Pred. No. 7.1e-39;
Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

Qy 28 AGLLLLPTFCQSGMENDTNNAKPTLPKTFRGAPPNSFEFFPSALEGWTGATITVKI 87
Db 12 AAIIAASLSC-SGTIOQTNRSKGRSLGKVDGT-----SHVTG-KGVTVTVF 59
Qy 88 KCPEESASHLVKNATMGVLTSSLTSLIPAIYLVFVGVGPANAVTLWMLFRTSICT 147
Db 60 SVDEFSAS-----VLAKLTIVFLPIVYIVFAVGLPSNGMALVFLFRTKKHP 109
Qy 148 TVFY-TNLAIADFLCVTLFPFKIAYHLNGNNWVGEVLCRATTIVFYGNMYCSILLACI 206
Db 110 AVIYMANLADLLSVIWPFLPKIAYHIGNNWIYGEALCNVLGFFYRNMYCSILFMTCL 169
Qy 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQYLVQPDITTC 266
Db 170 SVORYWIVNPMGH-SRKKANIAIGISLAIWLLTLVTLPIYVVKOTIFIPALNITTC 228

QY 267 VHTCESSPPOLY-----YFISLAPFGFLIPFVLIIYCYAAIIRTLNA-----YDHR 314
 Db 229 V-----LPEQLLVGDMFNFLSLAIGVFLFAFLTASAYVLMIRLRSSAMDENSEK 281
 QY 315 WLWYKASLLILVIFTCFAPSNIILIIHANYNNYNTDG---LYFIYLIALCLGSLNSC 371
 Db 282 RKRAIKLIVVLGMYLICFTPSNLLLVH---YFLIKSQSQSHVYALYIIVALCLSTLNSC 338
 QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREOQGP 404
 Db 339 IDPFVYVYVSHDRDHAKNALLCRSVRTVKMQVP 373

RESULT 36
 US-08-476-976-63
 ; Sequence 63, Application US/08476976
 ; Patent No. 5874400
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,976
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-0763
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 63:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 397 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-476-976-63

Query Match 26.4%; Score 569.5; DB 2; Length 397;
 Best Local Similarity 34.2%; Pred. No. 7.1e-39;
 Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

QY 28 AGLILLPLTFQSGMENDTNLAKPTLPKTFGAPPNSFEFPFSALEGTGATITVKI 87
 Db 12 AAILLAASLSC-SGTIGQTRSKGRSLICKVDGT-----SHVTG-KGVTVETVF 59
 QY 88 KCPEESASHLVKVNATMGYITSSLSKLIPIAYLLVFWGVGPANAVTLMMLFRTSIC 147
 Db 60 SVDEFAS-----VLAKGLTIVFLPIVYITVFAVGLPSNGMALVFLFRTKKHP 109
 QY 148 TVFY-TNLATADFLFCVTLFPKTAHUNGNNWFGVLCRATTVIFGNNYCSILLIACI 206
 Db 110 AVIYMANLALADLLSVIFWFLPKIAYHIGNNWYGEALCNVLIGFFFYRNMYCSILFMTCL 169

QY 207 SINRYLAIYVHPYTRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITICH 266
 Db 170 SVORYWVIVNPMGH-SRKKANIAIGISLAIMLLTLVITPIYVVKQTIPIPALNITICH 228
 QY 267 VHTCESSPPOLY-----YFISLAPFGFLIPFVLIIYCYAAIIRTLNA-----YDHR 314
 Db 229 V-----LPEQLLVGDMFNFLSLAIGVFLFAFLTASAYVLMIRLRSSAMDENSEK 281
 QY 315 WLWYKASLLILVIFTCFAPSNIILIIHANYNNYNTDG---LYFIYLIALCLGSLNSC 371
 Db 282 RKRAIKLIVVLGMYLICFTPSNLLLVH---YFLIKSQSQSHVYALYIIVALCLSTLNSC 338
 QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREOQGP 404
 Db 339 IDPFVYVYVSHDRDHAKNALLCRSVRTVKMQVP 373

RESULT 37
 US-08-474-410-63
 ; Sequence 63, Application US/08474410
 ; Patent No. 6043212
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,410
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 63:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 397 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-474-410-63

Query Match 26.4%; Score 569.5; DB 3; Length 397;
 Best Local Similarity 34.2%; Pred. No. 7.1e-39;
 Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

QY 28 AGLILLPLTFQSGMENDTNLAKPTLPKTFGAPPNSFEFPFSALEGTGATITVKI 87
 Db 12 AAILLAASLSC-SGTIGQTRSKGRSLICKVDGT-----SHVTG-KGVTVETVF 59
 QY 88 KCPEESASHLVKVNATMGYITSSLSKLIPIAYLLVFWGVGPANAVTLMMLFRTSIC 147

Db 60 SVDFSGAS-----VLAKLTITVFLPIVTVFAVGLPSNGMALWFLFRTKKKHP 109
QY 148 TVFY-TNLADEFLFCVTLFPFKIAYHLNGNNWVGEVLCRATTVIYGVNMYCSILLACI 206
Db 110 AVIYMANLAADLWSVIFPKIAYHGNWNIYGEALCNVLGFFYRNMYCSILFMTCL 169
QY 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQPDITTC 266
Db 170 SVQRYVWVNMGMH-SRKANIALIGISLAIWLTLTLLVPLVYVKQTIPIPALNITTC 228
QY 267 VHTCESSPFQY-----YFISLAPFFGLIPFVLIYYCYAAIIRTLNA-----YDHR 314
Db 229 V-----LPEQLLVGMFNYFLSLAIGVFLFPAFLTASAYVLMIRLSSAMSENSEKK 281
QY 315 WLVVVKASLLIVFTICFAPSNIILIIHHANYNNYNTDG---LYFIYLIALCLGSLNSC 371
Db 282 RKRAIKLIVTLVGLMYLICFTPSNLLLVH---YFLIKSQCSQSHVYALIVLALCLSTLNSC 338
QY 372 LDPLFLYPLMSKT-RNHS-TAYLTKXXNDLREQQP 404
Db 339 IDPFVYFVSHDFRDHAKNALLCRESVRTVKQMOPV 373

RESULT 38

US-08-097-938-7
; Sequence 7, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-097-938-7

Query Match 24.4%; Score 526; DB 1; Length 425;
Best Local Similarity 32.9%; Pred. No. 2.6e-35;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY 14 GTQVIMKALIFAAGLILLLPFCQSGMENDTNLAKPTLPKTRGAPPN-SPEFFP 72
Db 2 GPRLLVAACFLSCGLFLLSARTRARPESKATN----ATLDRSLRNPNDKPEFFWE 57
QY 73 SALEGWTGAT----ITVKIKCPESASHLVKNATMGYLTSSLSLTKLIPAILLVFVVG 128

Db 58 DEEKESGLTEYRLVINSKSSPLQKQPAFISEDASGYLTSSWLTLPVPSVVTGVFWSL 117
QY 129 PAN--AVTLWMLFFTRISICTIVFTYNLAIDFLCVTLFPFKIAYHLNGNNWVGEVLCR 186
Db 118 PLNIMAVIVFLMKVKK-PAVVYMLHLATADVFLSVLPFKISYFSGSQFGSELGR 176
QY 187 ATTIVFYGNMYCSILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
Db 177 FTAAFYCNWYASILLMTVISIDRFVAVYPMQSLSWRTLGRASF---TCLAIWALAIAAG 233
QY 244 MLFPFLIKQBYLVQPDITTCCHDVHTCESSPFQYFISLAPFFGLIPFVLIYYCYAA 303
Db 234 VVPLVLKEQTIQVPGNITTCCHDVLTLEGL-YVAYYFSAFSAVFFFPVPLIISTVCYVS 292
QY 304 IIRTLN-----AYDRWLWTVVKASLLILVITFCFAPSNIILIIHHANY-YNNYNTDG 355
Db 293 IIRCLSSAVANRKSRAFL---LSAAVFCIFIIICFGPTNVLLIAHYSFLSHTSTTAA 349
QY 356 YFIYLIALCLGSLNSCLDPFLYFLMS 381
Db 350 YFAYLLVCVSSISSCIDPLIYYAS 375

RESULT 39

US-08-476-000-7
; Sequence 7, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-476-000-7

Query Match 24.4%; Score 526; DB 1; Length 425;
Best Local Similarity 32.9%; Pred. No. 2.6e-35;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

QY 14 GTQVIMKALIPAAAGLLLLPTFCOSGMENDTNNAKPTLPKIKTRGAPPN-SPEEFPF 72
 Db 2 GPRRLLLVAACFSLCGFLLSARTRARRPESKATN---ATLDRSFLLRNPNDKYEPFWE 57
 QY 73 SALEGTGAT---ITVKIKCPESASHLVKNATMGYLTSSLSSTKLIPAIYLLVFWGV 128
 Db 58 DEKNESGLTEYRLVINSKSSPQKQLPAFISDASGYLTSSWLTFLVFSVYTGFWVSL 117
 QY 129 PAN--AVTLMMLFFRTRISCTTVFYTNLAIADEFLFCVTLPFKIAYHLNGNWNWVFGVLCR 186
 Db 118 PLNIMAIVFILKMKVKK-PAVVYMLHLATADLVFVSVPFKISYVFGSDMQFGSELGR 176
 QY 187 ATTIVFYGNMYCSILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
 Db 177 FVTAAFYCNMYASILLMTVISIDRFLAVVPMQSLSWRTLGRASF---TCLAIWALAIAAG 233
 QY 244 MLPFFILKQEVYLVQPDITTCCHDVHNTCESSSPFQLYYFISLAFPGFLIPFVLIYYCYAA 303
 Db 234 VPLVLUKEQTIQVPLNITTCCHDVNLNETLLEG-YAYYFSAFSAVFFVPLIISTVCYVS 292
 QY 304 IIRTLN-----AYDHRWLWYVKASLLILVIPTICFAPSNIILIIHHANY-YNNNTDGL 355
 Db 293 IIRCLSSAVANRSKSRALF---LSAAVFCIIICFGPTNVLIIAHYFSLSHTSTTEAA 349
 QY 356 YFIYLIACLSGNSCLDPLFLYFMS 381
 Db 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 40

US-08-472-840-7
 ; Sequence 7, Application US/08472840
 ; Patent No. 5763575
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/472,840
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 425 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-472-840-7

Query Match 24.4%; Score 526; DB 1; Length 425;
 Best Local Similarity 32.9%; Pred. No. 2.6e-35;
 Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
 QY 14 GTQVIMKALIPAAAGLLLLPTFCOSGMENDTNNAKPTLPKIKTRGAPPN-SPEEFPF 72
 Db 2 GPRRLLLVAACFSLCGFLLSARTRARRPESKATN---ATLDRSFLLRNPNDKYEPFWE 57
 QY 73 SALEGTGAT---ITVKIKCPESASHLVKNATMGYLTSSLSSTKLIPAIYLLVFWGV 128
 Db 58 DEKNESGLTEYRLVINSKSSPQKQLPAFISDASGYLTSSWLTFLVFSVYTGFWVSL 117
 QY 129 PAN--AVTLMMLFFRTRISCTTVFYTNLAIADEFLFCVTLPFKIAYHLNGNWNWVFGVLCR 186
 Db 118 PLNIMAIVFILKMKVKK-PAVVYMLHLATADLVFVSVPFKISYVFGSDMQFGSELGR 176
 QY 187 ATTIVFYGNMYCSILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
 Db 177 FVTAAFYCNMYASILLMTVISIDRFLAVVPMQSLSWRTLGRASF---TCLAIWALAIAAG 233
 QY 244 MLPFFILKQEVYLVQPDITTCCHDVHNTCESSSPFQLYYFISLAFPGFLIPFVLIYYCYAA 303
 Db 234 VPLVLUKEQTIQVPLNITTCCHDVNLNETLLEG-YAYYFSAFSAVFFVPLIISTVCYVS 292
 QY 304 IIRTLN-----AYDHRWLWYVKASLLILVIPTICFAPSNIILIIHHANY-YNNNTDGL 355
 Db 293 IIRCLSSAVANRSKSRALF---LSAAVFCIIICFGPTNVLIIAHYFSLSHTSTTEAA 349
 QY 356 YFIYLIACLSGNSCLDPLFLYFMS 381
 Db 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

Search completed: June 24, 2003, 12:07:21
 Job time : 21.0221 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:52:41 ; Search time 52.0638 Seconds
(without alignments)
1044.224 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIM.....AVLTXXNDLREQGPQSORT 408

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries.

Database :

A Genesecq_101002.*

1: /SID22/gcgdata/genesecq/genesecp-emb1/AA1980.DAT.*
2: /SID22/gcgdata/genesecq/genesecp-emb1/AA1981.DAT.*
3: /SID22/gcgdata/genesecq/genesecp-emb1/AA1982.DAT.*
4: /SID22/gcgdata/genesecq/genesecp-emb1/AA1983.DAT.*
5: /SID22/gcgdata/genesecq/genesecp-emb1/AA1984.DAT.*
6: /SID22/gcgdata/genesecq/genesecp-emb1/AA1985.DAT.*
7: /SID22/gcgdata/genesecq/genesecp-emb1/AA1986.DAT.*
8: /SID22/gcgdata/genesecq/genesecp-emb1/AA1987.DAT.*
9: /SID22/gcgdata/genesecq/genesecp-emb1/AA1988.DAT.*
10: /SID22/gcgdata/genesecq/genesecp-emb1/AA1989.DAT.*
11: /SID22/gcgdata/genesecq/genesecp-emb1/AA1990.DAT.*
12: /SID22/gcgdata/genesecq/genesecp-emb1/AA1991.DAT.*
13: /SID22/gcgdata/genesecq/genesecp-emb1/AA1992.DAT.*
14: /SID22/gcgdata/genesecq/genesecp-emb1/AA1993.DAT.*
15: /SID22/gcgdata/genesecq/genesecp-emb1/AA1994.DAT.*
16: /SID22/gcgdata/genesecq/genesecp-emb1/AA1995.DAT.*
17: /SID22/gcgdata/genesecq/genesecp-emb1/AA1996.DAT.*
18: /SID22/gcgdata/genesecq/genesecp-emb1/AA1997.DAT.*
19: /SID22/gcgdata/genesecq/genesecp-emb1/AA1998.DAT.*
20: /SID22/gcgdata/genesecq/genesecp-emb1/AA1999.DAT.*
21: /SID22/gcgdata/genesecq/genesecp-emb1/AA2000.DAT.*
22: /SID22/gcgdata/genesecq/genesecp-emb1/AA2001.DAT.*
23: /SID22/gcgdata/genesecq/genesecp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	Query Match	Length	ID	Description
1	1998	92.6	374	19 AAW51406	Human protease-act
2	1435.5	66.1	369	19 AAW51405	Mouse protease-act
3	606.5	28.1	399	16 AAR66922	Murine C140 recept
4	606.5	28.1	399	17 AAW01954	Murine C140 recept
5	600.5	27.8	395	16 AAR66920	Murine C140 recept
6	600.5	27.8	395	17 AAW01952	Murine C140 recept
7	589.5	27.3	397	21 AAB35641	Human PAR-2 protei
8	585.5	27.1	398	16 AAR66921	Human C140 recept
9	585.5	27.1	398	17 AAW01953	Human C140 recept
10	583	27.0	394	19 AAW51408	Human protease-act

11	569.5	26.4	397	16 AAR66923	Human C140 recept
12	569.5	26.4	397	17 AAW01955	Human C140 recept
13	531.5	24.6	396	20 AAY15081	Mouse protease-act
14	525	24.3	408	20 AAY50137	FLAG epitope-tagge
15	523.5	24.3	385	20 AAY50139	Mutant human prote
16	523	24.2	425	13 AAR27240	Human thrombin rec
17	523	24.2	425	19 AAW51407	Human protease-act
18	523	24.2	425	20 AAY49570	Human thrombin rec
19	523	24.2	425	23 AAG80697	Human thrombin pro
20	523	24.2	425	23 AAE17032	Human thrombin pro
21	523	24.2	426	21 AAY45035	Human thrombin rec
22	521.5	24.2	385	20 AAY15082	Human protease-act
23	521.5	24.2	385	20 AAY50135	Human protease-act
24	521.5	24.2	385	21 AAY45036	Human protease-act
25	521.5	24.2	385	22 AAB47623	Human PAR4. Homo
26	521	24.2	385	20 AAY50138	Mutant human prote
27	519	24.1	425	23 AAG35300	Human PAR1 type th
28	516.5	23.9	425	15 AAR60698	Fragment of the hu
29	512	23.7	892	18 AAW16314	Human thrombin rec
30	507.5	23.5	402	23 AAG35298	Human PAR1 type th
31	499.5	23.2	371	23 AAG35299	Human PAR1 type th
32	471.5	21.9	359	19 AAG69598	Mouse G-protein co
33	461	21.4	359	21 AAY71304	Human orphan G pro
34	461	21.4	359	21 AAB02838	Human G-protein co
35	460	21.3	359	21 AAY84815	A human G-protein
36	460	21.3	359	21 AAY69485	Amino acid sequenc
37	460	21.3	359	22 AAG78530	G-protein coupled
38	460	21.3	359	22 AAG80966	Human NGPCRS #2.
39	460	21.3	359	22 AAB62285	Human G-protein co
40	432.5	20.1	374	22 AAE04390	Turkey P2Y nucleot
41	428	19.8	537	23 AAU74538	Human P2Y purinoco
42	413.5	19.2	339	21 AAY90652	Human mutant G pro
43	411	19.1	322	15 AAR48715	G-protein coupled
44	411	19.1	322	17 AAW02687	G-protein coupled
45	410.5	19.0	339	15 AAR53752	Seven transmembran

ALIGNMENTS

RESULT 1
AAW51406
ID AAW51406 standard; Protein; 374 AA.
XX
AC AAW51406;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 3 (PAR3).
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis.
XX
OS Homo sapiens.
XX
FH Key
FT Modified-site /note= "Asn is N-glycosylated"
FT Cleavage-site /note= "thrombin cleavage site"
FT Peptide /note= "hirudin-like sequence"
FT Modified-site /note= "Asn is N-glycosylated"
FT Domain /label= TM1
FT /note= "transmembrane domain 1"
FT Domain /label= TM2
FT /note= "transmembrane domain 2"

FT Domain 168..191
 FT /label= TM3
 FT /note= "transmembrane domain 3"
 FT Domain 207..231
 FT /label= TM4
 FT /note= "transmembrane domain 4"
 FT Domain 261..286
 FT /label= TM5
 FT /note= "transmembrane domain 5"
 FT Domain 301..323
 FT /label= TM6
 FT /note= "transmembrane domain 6"
 FT Modified-site 331..333
 FT /note= "Asn is N-glycosylated"
 FT Domain 335..360
 FT /label= TM7
 FT /note= "transmembrane domain 7"
 FT XX
 PN WO9818456-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 29-OCT-1997; 97WO-US19732.
 XX
 PR 30-OCT-1996; 96US-0742440.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Connolly A, Coughlin SR, Ishihara H;
 XX
 XX WPI; 1998-271905/24.
 DR N-PSDB; AAV07372.
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 41-42; 74pp; English.
 XX
 CC This polypeptide comprises human protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372), and shows homology to human PAR1
 CC and PAR2 (see AAW51407 and AAW51408). The mouse PAR3 amino acid
 CC sequence (see AAW51405) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC compounds for their ability to act as agonists or antagonists of
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,
 CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 374 AA;
 Query Match 92.6%; Score 1998; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.3e-198;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 MKALIFAAAGLLLLPTFCOSGMENDNNLAKPTLPKTRGAPPNSFEFFPSSALEGWT 79
 Db 1 MKALIFAAAGLLLLPTFCOSGMENDNNLAKPTLPKTRGAPPNSFEFFPSSALEGWT 60
 QY 80 GATITVKIKCPESASHLVKNAATMGYLTSSLSSTKLIPAIYLLVFGVGPANAVTLWMLF 139
 Db 61 GATITVKIKCPESASHLVKNAATMGYLTSSLSSTKLIPAIYLLVFGVGPANAVTLWMLF 120
 QY 140 FRTRISCTTVFYNTLAIADFLFCVTLPFKIAYHLNGNNWVFGVLCRATTVIFYGNMYS 199

Db 121 FRTRISCTTVFYNTLAIADFLFCVTLPFKIAYHLNGNNWVFGVLCRATTVIFYGNMYS 180
 QY 200 ILLACISINRYLAIVHPETRYGLPKHTYALATCGLVWATVLYMLPFFILKQEYVYVQ 259
 Db 181 ILLACISINRYLAIVHPETRYGLPKHTYALATCGLVWATVLYMLPFFILKQEYVYVQ 240
 QY 260 DITTCDDVHNTCESSPPFOLYYFISLAPFGFLIPFVLIICYAAIIRTINAYDHRWLWV 319
 Db 241 DITTCDDVHNTCESSPPFOLYYFISLAPFGFLIPFVLIICYAAIIRTINAYDHRWLWV 300
 QY 320 KASLLILVITTCFAPSNIILIIHHANYNNNTDGLYFYIYALCLGSLNSCLDPFLYFL 379
 Db 301 KASLLILVITTCFAPSNIILIIHHANYNNNTDGLYFYIYALCLGSLNSCLDPFLYFL 360
 QY 380 MSKTRNHSTAYLTK 393
 Db 361 MSKTRNHSTAYLTK 374
 RESULT 2
 AAW51405
 ID AAW51405 standard; Protein; 369, AA.
 XX
 AC AAW51405;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Mouse protease-activated receptor 3 (PAR3).
 XX
 KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis.
 XX
 OS Mus sp.
 XX
 PN WO9818456-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 29-OCT-1997; 97WO-US19732.
 XX
 PR 30-OCT-1996; 96US-0742440.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Connolly A, Coughlin SR, Ishihara H;
 XX
 XX WPI; 1998-271905/24.
 DR N-PSDB; AAV07372.
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 39-40; 74pp; English.
 XX
 CC This polypeptide comprises mouse protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372). The human PAR3 amino acid
 CC sequence (see AAW51406) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC compounds for their ability to act as agonists or antagonists of
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,

CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.
XX
SQ Sequence 369 AA;
Query Match 66.1%; Score 1425.5; DB 19; Length 369;
Best Local Similarity 71.6%; Pred. No. 3e-139;
Matches 260; Conservative 44; Mismatches 58; Indels 1; Gaps 1;
QY 20 MKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPNSFEFFPFALEGGT 79
DB 1 MKLILVAAGLLFLPVTVCQSGI-NVSDNSAKPTLTIKSPNGPQNTFEFFPLSDIEGWT 59
QY 80 GAITIVKIKCPESASHLVKNATMGVLTSSLTSLKLPALYLLVVFVGVVPANAVTLWLF 139
DB 60 GATTTIKAECPEDSISTLVNNATIGYLRSSLSQVIPAIIYLLVVGVPESNTVTLWKL 119
QY 140 FRTRISCTTVFYTNLAIDFLFCVTLFPFKIAYHLGNWVFGVLCRATTIVFYGNMYS 199
DB 120 LRTKSLVIFHNLAIDFLFCVTLFPFKIAYHLGNWVFGVLCRATTIVFYGNMYS 179
QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFFLKQEYLVQP 259
DB 180 ILILTGMGINRYLATAHPFTYQKLPKSPSLMCGIWMVFLYMLPFFVLKQEXHLVHS 239
QY 260 DIITCHDVNTCESSSPFQLYYFISLAPFGFLIPFVLIIVCYAAIIRTNAYDHRMLWYV 319
DB 240 EITTCDDVDACESPSSEFYFVSFAFFGFLIPFVIIIFCYTTLIHKLKSKDRIWLYI 299
QY 320 KASILLIVITICFAPSNIILIIHHANYNNYNTDGLFYLIALLCLGSLNSCLDPFLYFL 379
DB 300 KAVLLILVITICFAPNTIILVIHHANYNNYNTDGLFYLIALLCLGSLNSCLDPFLYFL 359
QY 380 MSK 382
DB 360 MSK 362
RESULT 3
AAW01954
ID AAW01954 standard; Protein; 399 AA.
XX
AC AAW01954;
XX
XX
XX 22-AUG-1995 (first entry)
XX Murine C140 receptor deduced from cDNA.
XX G-protein-coupled receptor; G-protein; C140 receptor.
XX Mus musculus.
XX WO9503318-A.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US08536.
XX
XX 26-JUL-1993; 93US-0097938.
XX (CORT-) COR THERAPEUTICS.
XX Scarborough RM, Sundelin J;
XX WPI; 1995-075182/10.
XX N-PSDB; AAQ84559.
XX New DNA encoding recombinant C140 receptor - and novel agonists
PT and antagonists and specific antibodies with therapeutic and
PT diagnostic applications.
XX
XX Example; Fig 10; 57pp; English.
XX
XX A cDNA library from a mouse stomach was constructed in lambda gt10

CC and screened with a probe encompassing the C140 genomic clone (see
CC AAQ84557). A single phage clone was isolated and cut with EcoRI. The
CC insert was cloned in pBluescript and pSG5 and sequenced. The
CC complete nt sequence and deduced AA sequence is given in AAQ84559 &
CC AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
CC end. The 5' end of the apparent coding region differs from the 5'
CC end of the ORF of genomic DNA; it is believed that the 5' end of
CC the cDNA sequence is correct.
XX
SQ Sequence 399 AA;
Query Match 28.1%; Score 606.5; DB 16; Length 399;
Best Local Similarity 35.4%; Pred. No. 3.1e-54;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;
QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPFALEGGT 79
DB 5 SLAWLLGGITLLAASVCSRTENLAPGRNNSKRSLSIGRLETQPP-----IT 51
QY 80 GAITIVK--IKCPESASHLVKNATMGVLTSSLTSLKLPALYLLVVFVGVVPANAVTLWM 137
DB 52 KGKVPVEPGFSIDFSAS-----ILTGKLTTFVLPVVVYIIVFVIGLPSNGMALWI 101
QY 138 LFRTRISCTTVFY-TNLAIDFLFCVTLFPFKIAYHLGNWVFGVLCRATTIVFYGNM 196
DB 102 FLFRTKKKHPAVIYMANLALADLLSVIWPFLKISYHLGNWVYGEALCKVLIGFFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL--VWATVFLYMLPFFFLKOE 253
DB 162 YCSILFMTCLSVQRYWVIVNPM---GHPKK-ANIAVGSLALWLLIFLVTPLTYVMKQT 217
QY 254 YVLVQPDITTCDDVNTCESSSPFQLY-----YFISLAPFGFLIPFVLIIVCYAAIIR 307
DB 218 IYIPALNITTCDDV-----LPEEVLGDMENYFLSLAIGVFLFPALLTASAYVLMKT 270
QY 308 LNA-----YDHRWLMYVKASLLILVITICFAPSNIILIIHHANYNNYNTDGLFYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNIILVHVHFLIKTQRSHVYALYV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
DB 331 ALCLSLNSCIDPFVYFVSKDFRDA 357
RESULT 4
AAW01954
ID AAW01954 standard; Protein; 399 AA.
XX
AC AAW01954;
XX
XX 02-APR-1997 (first entry)
XX Murine C140 receptor.
XX
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX Mus sp.
XX
XX Key Location/Qualifiers
FH Peptide 1..31
FT /note= "the signal peptide differs from that
FT encoded by a genomic DNA sequence for
FT this receptor (see AAW01952), the signal
FT sequence given here is believed to be
FT the correct sequence"
FT Protein 32..399
FT /note= "mature protein"
XX WO9623225-A1.
XX
XX 01-AUG-1996.
XX
XX

Db 17 VMLLLRFLCTGRNNSKGRSLIGLETOPPTGKGVPEPGFSIDF----- 62
Qy 85 VKIKPEESASHLHVKNATMGYLTSLSTKLIPAIYLLVGVGVPANAVTLMLFFRTRS 144
Db 63 -----SAS-----ILTGKLTTFVLPVWYIIVFVIGLPSNGMALWFLFRTKK 104
Qy 145 ICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNVFGVLCRATTVIFGNNYCSILL 203
Db 105 KHPAVIYMANLALADLLSVIFPKISYHLGNNVYGEALCKVLIGFFGNNYCSILFM 164
Qy 204 ACISINRYLAIVHPFTYRGPKHYALVTCGL---VMATVFLYMLPFILKQEVYLVQPD 260
Db 165 TCLSQRVYVIVNPM---GHPKPK-ANIAVGSVLAIWLLIFLVTIPYVWKQIYIPALN 220
Qy 261 ITTCHDVHNTCESSPPQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA---- 310
Db 221 ITTCHDV-----LPEEVLVGMDFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMD 273
Qy 311 --YDHRWLWTVKASLLILVITTCFAPSNIILIIHHANYNNYNTDGLFYLIYALCLGSL 368
Db 274 EHSNKRQRAIRLIITVLAIFYCFRPSNLLLVVHYFLIKTRQSHVYALYVALCLSTL 333
Qy 369 NSCLDPFLYFLMSKT-RNHS 387
Db 334 NSCIDPFVYFVSKDFRDHA 353

RESULT 6
AAW01952
ID AAW01952 standard; Protein; 395 AA.
AC AAW01952;
DT 01-APR-1997 (first entry)
XX Murine C140 receptor, including putative signal sequence.
DE C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX Mus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /note= "putative signal peptide, differs from
FT signal peptide encoded by a cDNA clone of
FT this receptor (see AAW01954), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT
FT Protein 28..395
FT /note= "mature protein"
FT
FT Modified-site 29
FT /note= "potential Asn-linked glycosylation site"
FT
FT Cleavage-site 34..35
FT /note= "putative protease receptor cleavage site"
FT
FT Region 78..100
FT /note= "transmembrane region I"
FT
FT Region 108..128
FT /note= "transmembrane region II"
FT
FT Region 148..169
FT /note= "transmembrane region III"
FT
FT Region 188..210
FT /note= "transmembrane region IV"
FT
FT Modified-site 220
FT /note= "potential Asn-linked glycosylation site"
FT
FT Region 244..264
FT /note= "transmembrane region V"
FT
FT Region 286..306
FT /note= "transmembrane region VI"
FT
FT Region 324..345
FT /note= "transmembrane region VII"
PN W09623225-A1.

XX 01-AUG-1996.
XX PD
XX 25-JAN-1996; 96WO-US01179.
XX PF
XX 25-JAN-1995; 95US-0390301.
XX PR
XX (CORT-) COR THERAPEUTICS INC.
XX PA
XX Scarbrough RM, Sundelin J;
XX WPI; 1996-362813/36.
XX N-PSDB; AAT32036.
XX Vector for expression C140 cell surface receptor in host cell
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
XX
XX Example 1; Fig 1A-B; 60pp; English.
XX
CC AAW01952 represents the murine C140 receptor (C140R), including a
CC putative signal peptide (see features table). DNA encoding C140R may be
CC engineered so as to allow the recombinant expression of C140R in a
CC suitable host cell, i.e. by removing the native expression-control
CC sequences and replacing them with control sequences operable in the
CC host. Such a recombinant receptor can be expressed on the surface of
CC oocytes, this provides a good assay system for identifying
CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
CC receptor and a member of the "seven-pass" transmembrane receptor
CC superfamily (peptide chain of the receptor passes through the cell
CC membrane seven times, producing seven transmembrane regions within the
CC receptor molecule). The C140 receptor is involved in controlling blood
CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
CC signalling from this receptor, resulting in an increase in blood pressure
CC and are therefore useful in pharmaceuticals for the treatment of
CC hypotension (low blood pressure). Conversely agonists (see
CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
CC of hypertension (high blood pressure).
XX
SQ Sequence 395 AA;
Query Match 27.8%; Score 600.5; DB 17; Length 395;
Best Local Similarity 34.8%; Pred. No. 1.3e-53;
Matches 132; Conservative 69; Mismatches 115; Indels 63; Gaps 11;
Qy 30 LLLLLPTFCOSGMENDTNL--AKPTLPKTRGAPPN---SPEEPFSALEGTCATIT 84
Db 17 VMLLRFLCTGRNNSKGRSLIGRLETOPPTGKGVPEPGFSIDF----- 62
Qy 85 VKIKPEESASHLHVKNATMGYLTSLSTKLIPAIYLLVGVGVPANAVTLMLFFRTRS 144
Db 63 -----SAS-----ILTGKLTTFVLPVWYIIVFVIGLPSNGMALWFLFRTKK 104
Qy 145 ICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNNVFGVLCRATTVIFGNNYCSILL 203
Db 105 KHPAVIYMANLALADLLSVIFPKISYHLGNNVYGEALCKVLIGFFGNNYCSILFM 164
Qy 204 ACISINRYLAIVHPFTYRGPL--KHTYALVTCGLVWATVFLYMLPFILKQEVYLVQPD 261
Db 165 TCLSQRVYVIVNPM---GHPKPKANIAVGSVLAIWLLIFLVTIPYVWKQIYIPALN 221
Qy 262 ITTCHDVHNTCESSPPQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA---- 310
Db 222 ITTCHDV-----LPEEVLVGMDFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMD 274
Qy 311 --YDHRWLWTVKASLLILVITTCFAPSNIILIIHHANYNNYNTDGLFYLIYALCLGSL 369
Db 275 EHSNKRQRAIRLIITVLAIFYCFRPSNLLLVVHYFLIKTRQSHVYALYVALCLSTL 334
Qy 370 NSCLDPFLYFLMSKT-RNHS 387
Db 335 NSCIDPFVYFVSKDFRDHA 353

```
RESULT 7
AAB35641
ID AAB35641 standard; Protein; 397 AA.
XX AC AAB35641;
XX DT 19-FEB-2001 (first entry)
XX DE Human PAR-2 protein.
XX KW PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
XX KW asthma; chronic obstructive pulmonary; arthritis;
XX KW inflammatory bowel; psoriasis; eczema; multiple sclerosis.
XX OS Homo sapiens.
XX PN WO200063371-A1.
XX PD 26-OCT-2000.
XX PF 17-APR-2000; 2000WO-GB01455.
XX PR 15-APR-1999; 99GB-0008513.
XX PA (UYSO-) UNIV SOUTHAMPTON.
XX PI Walls AF, Palmer K, Compton SJ, Cairns JA, Gough AC;
XX WP1; 2000-679599/66.
XX PT Protease activated receptor 2 variants useful for treating inflammatory
XX PT diseases such as asthma, arthritis and psoriasis, and as hypertensives,
XX PT has reduced sensitivity to trypsin -
XX PS Claim 2; Page 55; 59pp; English.
XX CC The present invention relates to a variant protease activated
XX CC receptor 2 (PAR-2). The invention is useful for identifying an
XX CC individual having a polymorphism in the ECL-2 region of one or
XX CC both PAR-2 gene alleles. The invention may be used to develop treatments
XX CC for inflammatory diseases such as asthma, chronic obstructive
XX CC pulmonary diseases, arthritis, inflammatory bowel diseases, psoriasis
XX CC and eczema, multiple sclerosis and to raise blood pressure.
XX SQ Sequence 397 AA;
Query Match 27.3%; Score 589.5; DB 21; Length 397;
Best Local Similarity 35.0%; Pred. No. 1.8e-52;
Matches 137; Conservative 65; Mismatches 138; Indels 51; Gaps 13;
QY 28 'AGLLLLLPTFCQSGMENDTNLAKPTLPKTFRGAPNSPEEPFSALEGTGATITVKI 87
DB 12 AAILLAASLSC-SGTIQGTNRSGRSLICKVDGT-----SHVTG-KGVTVETVF 59
QY 88 KCPESASHLHVNATGYLTSSLTKLIPAILLVFVGVGPANAVTLWMLFFRTSICT 147
DB 60 SVDEFSS-----VLTKLTTFIVTIVFVGLPSGNHALWFLRTKKHP 109
QY 148 TVFY-TNLAIADFLFCVTLFPKIAHLNGNNWVGEVLCRATTIVFYGNMYCSILLACI 206
DB 110 AVIYMANLADLLSVIWFPLKIAHYHGNWYGEALCNVLGFFPYGNMYCSILEFTCL 169
QY 207 SINRYLAIVHPFYRGIPKHYALVTCGLWATVFLWMLPFFILKQEYVLVQPDITCHD 266
DB 170 SVORYWIVNPMGH-SRKKANIAIGISLAIWLILVLTIPLYVVKQTFIPALNITCDD 228
QY 267 VHNTESSSPFQLY-----YFISLAPFFGLIPFLVLIYCAAIIRFLNA-----YDHR 314
DB 229 V-----LPEQLLVGMFNYFSLAIGVFLPAFLTASAYVLMIKRLSSANDENSEKK 281
QY 315 WLWYVKASLLILVITFCAPSNIILIHANYNNYNTDG---LYFTYLIALCLGSLNSC 371
DB 282 RKRAIKLIVTVLAMYLCITPSPNLLLVH---YFLIKSQSQSHVYALYIVALCLSTLNSC 338
QY 372 LDPFLYELMSKT-RNHS-TAYLTKKXNDLRE 400
DB 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 369
RESULT 8
AAR66921
ID AAR66921 standard; Protein; 398 AA.
XX AC AAR66921;
XX DT 22-AUG-1995 (first entry)
XX DE Human C140 receptor.
XX KW G-protein-coupled receptor; G-protein; C140 receptor.
XX OS Homo sapiens.
XX FH Key
XX FT Protein
XX FT Modified-site 31
XX FT Modified-site 223
XX FT Cleavage-site 37..38
XX FT Region 81..103
XX FT Region 11..132
XX FT Region 150..174
XX FT Region 191..212
XX FT Region 245..267
XX FT Region 289..309
XX FT Region 327..348
XX PN WO9503318-A.
XX QY 02-FEB-1995.
XX PF 26-JUL-1994; 94WO-US08536.
XX PR 26-JUL-1993; 93US-0097938.
XX PA (CORF-) COR THERAPEUTICS.
XX PI Scarborough RM, Sundelin J;
XX WP1; 1995-075182/10.
XX N-PSDB; AAQ84558.
XX New DNA encoding recombinant C140 receptor - and novel agonists
XX PT and antagonists and specific antibodies with therapeutic and
XX PT diagnostic applications.
XX PS Disclosure; Fig 2; 57pp; English.
XX CC The availability of genomic DNA encoding the mouse protease
XX CC C140 receptor (see Q84557) permitted the retrieval of the corresp.
XX CC human gene. A human genomic library cloned in the vector EMBL3 was
XX CC screened using the entire coding region of the murine clone as a
XX CC probe. The recovered human gene including the DNA sequence and the
XX CC deduced AA sequence are shown in Q84558 & R66921. Subsequent
XX CC experiments indicated that the human C140 gene is located in the
```


QY 207 SINRYLAIHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTCDD 266
 Db 170 SVQRYWIVNPMGH-SRKKANIAIGISLAIWLTLLVTLPIYVVKQTIFIPALNITTCDD 228
 QY 267 VHTNCSSPPOLY-----YFISLAPFGFLIPFVLIYCYAAIIRTLNA-----YDHR 314
 Db 229 V-----LPEQLLVGDMFNFLSLAIGVFLPFAFLTASAYVLMIRLSSAMDENSEKK 281
 QY 315 WLWYKASLLILVIFTCFAPSNIILIHANNYYNNTDG---LYFIYLIALCLGSLNSC 371
 Db 282 RKRAIKLIVTVLGMVLYICTPNSLLLVH---YFLIKSQSQSHVYALYIVALCLSTLNSC 338
 QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREQQOP 404
 Db 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMOP 373

RESULT 12

AAW01955
 ID AAW01955 standard; Protein; 397 AA.
 AC AAW01955;

02-APR-1997 (first entry)

Human C140 receptor.

XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
 KW antagonist; hypertension; hypotension; blood pressure.
 XX

OS Homo sapiens.

Key Location/Qualifiers

FT Peptide 1..27
 /note= "the signal peptide differs from that encoded by a genomic DNA sequence for this receptor (see AAW01953), the signal sequence given here is believed to be the correct sequence"

FT Protein 28..397
 /note= "mature protein"

PN WO9623225-A1.

XX 01-AUG-1996.

XX 25-JAN-1996; 96WO-US01179.

XX 25-JAN-1995; 95US-0390301.

XX (CORT-) COR THERAPEUTICS INC.

XX Scarborough RM, Sundelin J;

XX WPI; 1996-362813/36.

DR N-PSDB; AAT32039.

XX Vector for expression C140 cell surface receptor in host cell
 PT useful to identify C140 agonist and antagonists, which are
 PT antihypertensives and elevators of blood pressure, respectively

PS Example 5; Fig 11A-B; 60pp; English.

XX AAW01955 represents the human C140 receptor (C140R). DNA encoding C140R
 CC may be engineered so as to allow the recombinant expression of C140R in
 CC a suitable host cell, i.e. by removing the native expression-control
 CC sequences and replacing them with control sequences operable in the
 CC host. Such a recombinant receptor can be expressed on the surface of
 CC oocytes, this provides a good assay system for identifying
 CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
 CC receptor and a member of the "seven-pass" transmembrane receptor
 CC superfamily (peptide chain of the receptor passes through the cell
 CC membrane seven times, producing seven transmembrane regions within the

CC receptor molecule). The C140 receptor is involved in controlling blood
 CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
 CC signalling from this receptor, resulting in an increase in blood pressure
 CC and are therefore useful in pharmaceuticals for the treatment of
 CC hypotension (low blood pressure). Conversely agonists (see
 CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
 CC of hypertension (high blood pressure).

XX Sequence 397 AA;

Query Match 26.4%; Score 569.5; DB 17; Length 397;
 Best Local Similarity 34.2%; Pred. No. 2.1e-50;
 Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

QY 28 AGLLLLPTFCQSGMENDTNNAKPTLPKTRGAPPNSFEFFPFSALEGWTGATITVKI 87
 Db 12 AAILLAASLSC-SGTIOGTNRSSKGRSLIGKVDGT-----SHVTG-KGVTVETVF 59
 QY 88 KCPESASHLHVNATMGVLTSSLSKLIPIALYLLVGVGVANAVTLWMLPFRRSICT 147
 Db 60 SVDEFSSAS-----VLAKLTTVFLPIVYTVFAVGLFSGNMGALWVLFRTKKQHP 109
 QY 148 TVFY-TNLAIADFLFCVTLPEKIAYHLGNVNVFGEVLCRATTVIFYGNMYCSILLACI 206
 Db 110 AVIYMANLADLLSVIFWFLKIAIHGNNIYGEALCNVLIGFFYRMYCSILFTWCL 169
 QY 207 SINRYLAIHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTCDD 266
 Db 170 SVQRYWIVNPMGH-SRKKANIAIGISLAIWLTLLVTLPIYVVKQTIFIPALNITTCDD 228
 QY 267 VHTNCSSPPOLY-----YFISLAPFGFLIPFVLIYCYAAIIRTLNA-----YDHR 314
 Db 229 V-----LPEQLLVGDMFNFLSLAIGVFLPFAFLTASAYVLMIRLSSAMDENSEKK 281
 QY 315 WLWYKASLLILVIFTCFAPSNIILIHANNYYNNTDG---LYFIYLIALCLGSLNSC 371
 Db 282 RKRAIKLIVTVLGMVLYICTPNSLLLVH---YFLIKSQSQSHVYALYIVALCLSTLNSC 338
 QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREQQOP 404
 Db 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMOP 373

RESULT 13

AAV15081
 ID AAV15081 standard; Protein; 396 AA.

XX AAV15081;

XX 25-JAN-2000 (first entry)

XX Mouse protease-activated receptor 4 (PAR4).

XX Mouse protease-activated receptor 4; PAR4; thrombin receptor;
 KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
 KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
 KW wound; blood coagulation; heart attack; stroke; inflammatory response;
 KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
 KW pulmonary inflammation; ARDS; adult respiratory distress syndrome.

OS Mus musculus.

XX Key Location/Qualifiers

XX Cleavage-site 59..60
 FT /label= Thrombin_cleavage_site
 FT Peptide 60..65
 FT /label= PAR4_activating_peptide
 FT Protein 1..396
 FT /label= PAR4_protein

FT /note= "The sequence is stated to be 397 amino acids
 long in the specification but is shown to have 396 aa
 in figure-2 as well as Sequence Listing"

XX

SQ Sequence 385 AA;
Query Match 24.1%; Score 523.5; DB 20; Length 385;
Best Local Similarity 33.4%; Pred. No. 1.2e-45;
Matches 123; Conservative 62; Mismatches 142; Indels 41; Gaps 6;
QY 29 GLLLLLP-----TFQSGMENDTNLAKPTLPKTPRGAPPNSFEFFPFS 73
DB 3 GRLLWPLVLGSLGGTQTPSVYDESGTGCGDDSTPSILPAP--RGYP-----50
QY 74 ALSEGTGATTVIKKPEBSASHLVKNATMGVLTSSLTSLKLPALYLLVVFVGVANAV 133
DB 51 ---GQVCANDSTLELPDSSAALL-----LGV-----VPTRLVPALYGLVVLGVPANGL 97
QY 134 TLAWLFFRTSICCTVFTYNLATADELFCVTLFPKTYHNLGNWVGEVLGRATTVIF 193
DB 98 ALWVLTQAPRLPSTMLMNLATADLLALPPIAYHLRGQRPFGAACLATAALY 157
QY 194 GNMYSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKOE 253
DB 158 GHMYSVLLLAAYSLDRLYALVHLPLRRLALGRRLALGLCMAAWLMAALALPLTLQRT 217
QY 254 YVLVQPDITTCVHDVHNTCESSFPQLYFISLAFGLFIPFVLIIVCYAAIIRTLNAYDH 313
DB 218 FRLARSDRVLCHDALPLDAQSHWQ-PAFTCLALLGCFPLPLAMLLCYGATLHTLAASGR 276
QY 314 RLMWYVKASLLIIVITFCAPSNIILIIHHANYNNYNTDGLYFIYLIACLGSLNSCLD 373
DB 277 RYGHAKRLTAVLASAVAFVPSNLLLLHYDPSAMGNLYGAYVPSLAULSTLNSCVD 336
QY 374 PFLYFLMS 381
DB 337 PFIYIYVS 344

RESULT 16

AAR27240
ID AAR27240 standard; Protein; 425 AA.
XX AAR27240;
AC AAR27240;
XX 21-MAY-1998 (first entry)
XX Human thrombin receptor (TR).
XX
XX Diagnosis; cardiovascular disease; wound healing; restenosis;
KW thrombosis; unstable angina treatment; myocardial infarction;
KW thrombotic; thromboembolytic stroke.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT peptide 1..26 /label= sig_peptide
FT peptide 28..425 /label= mat_protein
FT peptide 28..41 /note= "activation peptide RPEKATNATLDR"
FT cleavage_site 41..42 /notes "thrombin-catalysed cleavage-site"
FT disulfide_bond 175..254 /note= "analogous to rhodopsin and beta-2 adrenergic receptor"
FT domain 100..129 /notes "transmembrane domain I"
FT domain 137..160 /notes "transmembrane domain II"
FT domain 177..197 /note= "transmembrane domain III"
FT domain 219..239 /notes "transmembrane domain IV"
FT domain 266..294 /note= "transmembrane domain V"

FT domain 311..335 /note= "transmembrane domain VI"
FT domain 348..367 /note= "transmembrane domain VII"
FT region 34..36 /note= "consensus N-linked glycosylation site"
FT region 62..64 /note= "consensus N-linked glycosylation site"
FT region 75..77 /note= "consensus N-linked glycosylation site"
FT region 250..252 /note= "consensus N-linked glycosylation site"
FT region 259..261 /note= "consensus N-linked glycosylation site"
XX WO9214750-A.
XX 03-SEP-1992.
XX 19-FEB-1992; 92WO-US01312.
XX 19-FEB-1991; 91US-0657769.
XX 07-NOV-1991; 91US-0789184.
XX (CORT-) COR THERAPEUTICS INC.
XX (REGC) UNIV CALIFORNIA.
XX Coughlin SR, Scarborough RM;
XX WPI; 1992-316119/38.
XX N-PSDB; AAQ28568.
XX DNA encoding cell surface receptor for thrombin - useful for
XX determining thrombin in diagnosing e.g. cardiovascular diseases,
XX also to treat wound healing, restenosis etc.
XX Disclosure; Fig 1; 81pp; English.
XX The TR DNA sequence can be used in the prepn. of diagnostics to
XX determine thrombin levels in samples, and screening tools for
XX candidate substances which affect thrombin activity in vivo.
XX Thrombosis may be diagnosed in a mammal by measuring the presence,
XX absence or amt. of the cleaved activation peptide of the TR.
XX Sequence 425 AA;
Query Match 24.2%; Score 523; DB 13; Length 425;
Best Local Similarity 32.9%; Pred. No. 1.5e-45;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY 14 GTQVIKMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTPRGAPPN-SPEEPFF 72
DB 2 GPRLLLVAAACFSLCGPLLSARTRARRPESKATN----ATLDPKSLRNPNDKYEPFWE 57
QY 73 SALEGWTGAT---ITVKIKPEBSASHLVKNATMGVLTSSLTSLKLPALYLLVVFVGV 128
DB 58 DEEKNEGLTEYRLVSINKSSPLQKLPAFISDASGYLTSSWLTLPVPSVVTGVPVSL 117
QY 129 PAN--AVTLWMLFFRTRISCTTVFTYNLAIAFLPCVTLPPKTYHNLGNWVGEVLGR 186
DB 118 PLNIMAVVILKMKVKK-PAVVYMLHLATADLVFVSLPFPKISYIFSGDMQFSELGR 176
QY 187 ATTVIFYGNMYSILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
DB 177 FVTAAPFCNMYSILLMTVISIDRLAVVYPMQUSWRTLGRASP---TCLAIWALAIAG 233
QY 244 MLPFFILKOEYLVQPDITTCVHDVHNTCESSFPQLYFISLAFGLFIPFVLIIVCYAA 303
DB 234 VVPLVLKEQTIQVPGNLITTCVHDVHNTLEGG-YVAYFSAFSAFVFFPLIISTVCYVS 292
QY 304 IIRTLN-----AYDHRMLWYVKASLLILVITTCFAPSNIILIIHHANY-YNNYNTDGL 355
DB 293 IIRCLSSAVANRKKRALF---LSAAVFCIFICFGPTNVLIIAHYFSLSTTEAA 349

```

QY      356 YFIYLIALCGLNSCLDPEFLYFLMS 381
DB      177 FVTAAPFCNMYASILMTVISIDRFVAVYPMQSLSWRTLGRASF--TCLAIWALAIAAG 233

RESULT 17
AAW51407
ID      AAW51407 standard; Protein; 425 AA.
AC      AAW51407;
XX      XX
XX      Key Location/Qualifiers
XX      Binding-site 52..55
XX      /note= "thrombin binding site"
XX      Cleavage-site 41..42
XX      /note= "thrombin cleavage site"
XX      WO9818456-A1.
XX      07-MAY-1998.
XX      29-OCT-1997; 97WO-US19732.
XX      30-OCT-1996; 96US-0742440.
XX      (REGC ) UNIV CALIFORNIA.
XX      Connolly A, Coughlin SR, Ishihara H;
XX      WPI; 1998-271905/24.
XX      DNA encoding protease-activated receptor 3 - for detection of
XX      specific agonists and antagonists, potentially useful for treating
XX      e.g. thrombosis, atherosclerosis, inflammation etc.
XX      Example 1; Page 42-43; 74pp; English.
XX      This polypeptide comprises human protease-activated receptor 1
XX      (PAR1), a receptor that mediates thrombin signalling. The
XX      invention relates to novel mouse and human PAR3 (see AAW51405-06)
XX      that show homology to PAR1 and which are specific receptors for
XX      thrombin. They can be used to screen for specific agonists and
XX      antagonists of thrombin useful e.g. for treating atherosclerosis,
XX      thrombosis and inflammation.
XX      Sequence 425 AA;
XX      Query Match 24.2%; Score 523; DB 19; Length 425;
XX      Best Local Similarity 32.9%; Pred. No. 1.5e-45;
XX      Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

QY      14 GTQVIMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPN-SFEEFPF 72
DB      2 GPRLLLVACFSICGPLLARTARRPESKATN----ATLDRSFLLRNPNDKYEPFWE 57

QY      73 SALEGWGTAT-----ITVKIKCPESASHLHVKNATMGYLTSSISTKLIPAIYLLVFVGV 128
DB      58 DEEKNESGLTEYRLVRSINKSPLOKQLPAFISEDASGYLTSSWLTLPVPSVYGVTVFVSL 117

QY      129 PAN--AVTLWMVLPFRITSCTTYFYTNLAITADELFCVTLPFKIAYHLNGNNWVGEVLOR 186
DB      118 PLNIMAIVPILMKVKKK--PAVYVIMLHLATADLVFVSLVPFKISYVPSGSDWQFSELOR 176

QY      187 ATTIVFYGNMYCSILLACISINRYLAIVHP---FTYRGLPKHYALTVCGLVWATVFLY 243
DB      177 FVTAAPFCNMYASILMTVISIDRFVAVYPMQSLSWRTLGRASF--TCLAIWALAIAAG 233

QY      244 MLPFFILKQEYLVQPDITTCDDVHNTCESSPPFQLYYFISLAFFGLIPFFVLI11CYAA 303
DB      234 VVPLVLEQTIQVGLNITTCDDVHNTLLEG-YVAYYFSAFSAVFFVFLIISTVCYVS 292

QY      304 IIRTLN-----AYDHRWLWYVKASLLILVIFTTICFAPSNIILIIHANY-YNNNTDGL 355
DB      293 IIRCLSSAVANRKSRALF---LSAAVFCIFIIICFGPTNVLIIAHYFSLSHTSTTEA 349

QY      356 YFIYLIALCGLNSCLDPEFLYFLMS 381
DB      350 YFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 18
AA49570
ID      AA49570 standard; Protein; 425 AA.
XX      XX
XX      AC AA49570;
XX      XX
XX      DT 13-JAN-2000 (first entry)
XX      XX
XX      DE Human thrombin receptor protein sequence.
XX      XX
XX      KW Human; coding sequence polymorphism; vascular pathology gene;
XX      polymorphic site; phenotype correlation; forensic; paternity testing;
XX      medicine; genetic analysis; vascular disease.
XX      XX
XX      OS Homo sapiens.
XX      XX
XX      PN WO9950454-A2.
XX      XX
XX      PD 07-OCT-1999.
XX      XX
XX      PF 26-MAR-1999; 99WO-US06473.
XX      XX
XX      PR 01-APR-1998; 98US-0054272.
XX      XX
XX      PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX      XX
XX      PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;
XX      XX
XX      DR WPI; 1999-620066/53.
XX      DR N-PSDB; AA232191.
XX      XX
XX      PT Determination of polymorphisms in genes, especially those identifying
XX      predisposition to vascular disease
XX      PS Disclosure; Fig 35; 134pp; English.
XX      XX
XX      CC AA232159 to AA232194 represent reference alleles for specifically
XX      claimed nucleic acid sequences from the present invention which comprise
XX      polymorphic sites as given in a table in the specification, selected
XX      from 92 single nucleotide polymorphisms in which the nucleotide at the
XX      polymorphic site is different from a nucleotide at the same site in a
XX      reference allele. The nucleic acids, and primers and probes, are used to
XX      identify polymorphisms, which may predispose an individual to disease,
XX      especially a vascular disease. They can also be used in phenotype
XX      correlations, forensics, paternity testing, medicine or genetic
XX      analysis. AA49550 to AA49573 represent the proteins which correspond
XX      to some of the reference alleles.
XX      SQ Sequence 425 AA;
XX      Query Match 24.2%; Score 523; DB 20; Length 425;
XX      Best Local Similarity 32.9%; Pred. No. 1.5e-45;
XX      Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

QY      14 GTQVIMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPN-SFEEFPF 72
DB      118 PLNIMAIVPILMKVKKK--PAVYVIMLHLATADLVFVSLVPFKISYVPSGSDWQFSELOR 176

```


Db 293 IIRCLSSAVANKSKRALF---LSAAVFCFIICFGPTNVLLIAHYSLSTSTTEAA 349
QY 356 YFIYLIACLGSLNSCLDPFLYFLMS 381
Db 350 YFAYLLCVCVSSISICIDPLIYYAS 375
RESULT 20
AAE17032
ID AAE17032 standard; Protein; 425 AA.
AC AAE17032;
XX
DT 18-APR-2002 (first entry)
XX Human thrombin protein.
DE
KW Human; G protein-coupled receptor; GPCR; SLGP; cellular proliferation;
KW growth; differentiation; migration disorder; cancer; retinal; vasotropic;
KW optic disk neovascularisation; cytostatic; gene therapy; genetic disease;
KW tissue typing; forensic biology; antiarthritic; ischaemia; arthritis;
KW thrombin.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Domain 102..132
FT /label= TM1
FT /note= "Transmembrane domain"
FT Domain 138..165
FT /label= TM2
FT /note= "Transmembrane domain"
FT Domain 176..200
FT /label= TM3
FT /note= "Transmembrane domain"
FT Domain 215..241
FT /label= TM4
FT /note= "Transmembrane domain"
FT Domain 268..296
FT /label= TM5
FT /note= "Transmembrane domain"
FT Domain 313..340
FT /label= TM6
FT /note= "Transmembrane domain"
FT Domain 347..379
FT /label= TM7
FT /note= "Transmembrane domain"
XX W0200202602-A2.
XX
XX 10-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-US20751.
XX
XX 30-JUN-2000; 2000US-0608921.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Tsai F;
XX
XX WPI; 2002-140083/18.
XX
XX New isolated G protein-coupled receptor nucleic acid molecule, SLGP,
XX useful for diagnosing and treating cellular proliferation, growth,
XX differentiation, or migration disorders e.g. cancer, arthritis and
XX myocardial ischemia
XX
XX Disclosure; Page 158-159; 182pp; English.
XX
XX The present invention relates to an isolated G protein-coupled receptor
XX (GPCR) nucleic acid molecule, SLGP which encodes a polypeptide. SLGP is
XX useful as target or therapeutic agent for diagnosis and treatment of
XX cellular proliferation, growth, differentiation, or migration disorders

CC (e.g., cancer, arthritis, retinal and optic disk neovascularisation, and
CC tissue ischaemia, such as myocardial ischaemia). SLGP is useful in
CC screening assays, detection assays (e.g., chromosomal mapping, tissue
CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,
CC prognostic assays, monitoring clinical trials and pharmacogenomics), and
CC in methods of treatment (e.g., therapeutic and prophylactic). SLGP is
CC useful as reagents or targets in assays applicable to treatment and
CC diagnosis of SLGP-mediated or related disorders. SLGP is useful in gene
CC therapy, to express SLGP protein, to detect SLGP mRNA or a genetic
CC alteration in a SLGP gene, and to modulate SLGP activity. SLGP is useful
CC to map their respective genes on a chromosome, and thus locate gene
CC regions associated with genetic disease, to identify an individual from a
CC minute biological sample (tissue typing), and to aid in forensic
CC identification of a biological sample. The present sequence is human
CC thrombin protein used in the invention.
XX
SQ Sequence 425 AA;
Query Match 24.2%; Score 523; DB 23; Length 425;
Best Local Similarity 32.9%; Pred. No. 1.5e-45;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY 14 GTQVIKMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTFRGAPPN-SFEFFPF 72
Db 2 GPRRLLVAAFCFLGCLPLLSARTRARRPESKATN----ATLDRSLLRNPNPKYFFWE 57
QY 73 SALEGWGTAT---ITVKKICPEESASHLHVKNATMGYLTSSLTSLKLIPIAYLLVFGVGV 128
Db 58 DEEKNESGLTEYRLVYSINKSSPLQKLPAPFISEDASGYLTSSWLTFLVPSVYTGTVFVSL 117
QY 129 PAN--AVTLMLFFRTRISCTTVFYTNLAIADFLFCVTLFPFKIAYHLNGNNWVGEVLGR 186
Db 118 PLNIMAVVFLKMKVKK-PAVVYMLHLATADLVFVSLPFFKISYIFSGSDWQFGSELGR 176
QY 187 ATTIVFYGNMYCSILLACISINRYLIVHP---FTYRGLPKHYALVTCGLVNAITVFLY 243
Db 177 FVTAAPFCNMYSILLMTVISIDRFLAVVYPMQSLWSRTLGRASF--TCLAIWALAIAG 233
QY 244 MLPFFILKQEYLVQPDITTDHVNHTCESSPFQLYYFISLAFFGLFPFVLIICYAA 303
Db 234 VVPLVLKEQTIQVPGNLITTDHVNHTLLEG-YIAYYFSAFSAVFFVPLIISTVCYVS 292
QY 304 IIRTLN-----AYDHRWLWYVKASLLILVITFCFAPSNIILIIHHANY-YNNNTDGL 355
Db 293 IIRCLSSAVANKSKRALF---LSAAVFCFIICFGPTNVLLIAHYSLSTSTTEAA 349
QY 356 YFIYLIACLGSLNSCLDPFLYFLMS 381
Db 350 YFAYLLCVCVSSISICIDPLIYYAS 375
RESULT 21
AAE17032
ID AAY45035 standard; Protein; 426 AA.
XX
XX AAY45035;
XX
XX 31-MAY-2000 (first entry)
XX
XX Human thrombin receptor, a member of protease activated receptor family.
XX
XX Human; thrombin receptor; THR; PAR-1; protease activated receptor;
XX antiseptic molecule; PAR antibody; cytostatic; therapeutic;
XX metastatic tumour cell; placental implantation; invasive cell.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX 37..61
XX /note= "conserved unique region comprising protease
XX activated domain and hirudin binding domain."
XX Therapeutically useful antiseptic molecules are designed
XX based on this region".

FT	Peptide	42..51
FT	/label= TRAP	
FT	/note= "Thrombin-receptor activating peptide used	
FT	for producing anti-Thr antibodies"	
FT	Misc-difference	426
FT	/note= "Encoded by TAG"	
FT	Region	251..257
FT	/note= "conserved residues in second extracellular	
FT	loop used for designing therapeutically useful	
FT	antiseuse molecules "	
XX		
PX	WO200008150-A1.	
PN		
XX		
PD	17-FEB-2000.	
XX		
PP	05-FEB-1999;	99WO-IL00079.
XX		
PR	07-AUG-1998;	98IL-0125698.
XX		
PA	(HADA-) HADASIT MEDICAL RES SERVICES & DEV.	
XX		
PI	Bar-Shavit R;	
XX		
XX		
DR	WPI; 2000-205706/18.	
PS	N-PSDB; /AAZ50771.	
XX		
XX	Example 2; Fig 1b; 46pp; English.	
CC	The patent discloses a method to treat metastatic tumour cells using	
CC	an antiseuse molecule comprising a polynucleotide complementary to an	
CC	RNA sequence of a protease activated receptor (PAR) protein, or an	
CC	antibody capable of binding to a PAR protein. The antiseuse molecules and	
CC	antibodies of PAR protein are also used to treat disorders associated	
CC	with implantation of placenta. The present sequence is a	
CC	human thrombin receptor (THR) (also known as PAR-1), which is a	
CC	G-coupled protein belonging to the PAR family. This sequence is used to	
CC	produce antiseuse molecules or antibodies against PAR protein for	
XX	therapeutic use.	
XX		
SQ	Sequence	426 AA;
	Query Match	24.2%; Score 523; DB 21; Length 426;
	Best Local Similarity	32.9%; Pred. No. 1.5e-45;
	Matches	127; Conservative 76; Mismatches 153; Indels 30; Gaps 11
QY	14 GTQVIKMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTFEGAPPN-SPEEPFF 72	
Db	2 GPRULLVAACFSLGCPULSARTARRPESKATN----ATLDPRSFLLRNPDKYFPFWE 57	
QY	73 SALEGWGTGAT---ITVKIKCPESASHLVKNATMGYLTSLSSTKLIPAIYLLVFWVG 128	
Db	58 DEEKNEGLTEYRLVSINKSPLOQLPAFTISEDAGYLTSSMLTLFVPVSYTGVPVSL 117	
QY	129 PAN--AVTLWMLFRTSRICITVFVTNLIAIDFLFCVTLPFKTAIHLNGNNWFGEVLCR 186	
Db	118 PLNIMAIVFILKMVKK-PADVVMHLHATADVLFSVLFPKISYPFGSDMQFGSELCR 176	
QY	187 ATTVIFYGNMYCSILLACISINRYLATVHP---FTVRGLPKHTYALTCLGLWNTVFLY 243	
Db	177 FTAAPYCNMYASIIMTVISIDRELAVVYPMQSLSWRTLGRASF---TCLAIALAIAIG 233	
QY	244 MLFPFIKEYYLVOPDITTCCHDVHNTCESSPFLQYFIFSLAFPGFLIPFVLIICYAA 303	
Db	234 VVPLVLKEQTIQVPGNIITTCCHDLVNELLGS-YYAYYFAFSAVFFVFPVLIISTCYVS 292	
QY	304 IIRTLN-----AYDHRWLWYVKASLLILVIFTCFAPSNIILIHANY-YNNWDGL 355	
Db	293 IIRCISSSAVANRKKSRALF---LSAAVFCIFICFGPTNVLLIIAHYSFLSHSTTEAA 349	

[illegible]

Db 3 GRLLMLPLVLGFLSLSGGTQ-----TPSVYDESGSTGGDDSTPSILPAPRGYQGV----- 53
QY 87 IKCPESASHLHVKNATMGYLTSSLSKLIPIAYLLVVFVGV PANAVTLMLFFRTRIC 146
Db 54 --CANDSDT-LELPDSSRALLLGWVPTFLVPALYGLVVLVGLPANGALWVLTATQAPRLP 110
QY 147 TTVFYTNLAIDELFCVTLDFKTAHYHLNGNNWVGEVLCRAITVIFVGNMYCSILLIACI 206
Db 111 STMLLMNLATADULLALPPIAYHLRGORWPFGEAACRLATAALYGHMYGSLVLLAAV 170
QY 207 SINRYLAIVHPFTYRGPKHTYALVTCGLVWATVFLVWLPFFILKQBYLYLVOPDITTCCHD 266
Db 171 SLDRYLAIVLPLARALGRRLALGLCMAWLMWAAALALPLTQRTFLARSDRVLCCHD 230
QY 267 VHNCTSSPPQLYFYFISLAFPGFLIPFVLIYYCYAAIIRTNAYDHRWLWYVKASILIL 326
Db 231 ALPLDAQASHWQ-PAFTCALLGCFLLPALLMCLCYGATLTLAASGRRYGHARLRTAVL 299
QY 327 VITICFAPSNIILIIHANYNNYNTDGLFYIYIALCLGSLNSCLDPPFLYFLMS 381
Db 290 ASAVAFVPSNLLLLHYSDSPSPSAGNLYGAYVPSIALSTLNSCVDPFYIYYVS 344

RESULT 23
AA50135
ID AA50135 standard; Protein; 385 AA.
XX AC AA50135;
XX DT 31-JAN-2000 (first entry)
XX DE Human protease-activated receptor PAR4.
XX KW Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;
KW platelet; proliferation; differentiation; mediation;
KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW growth factor; production.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Region 1..78 /note= "Extracellular N-terminal region"
FT Peptide 1..17 /note= "Signal peptide"
FT Cleavage-site 17..18 /note= "Cleaved by signal peptidase"
FT Protein 18..385 /note= "Mature non-activated human PAR4"
FT Cleavage-site 47..48 /note= "Cleaved by thrombin or trypsin to activate"
FT Protein 48..385 /note= "Activated human PAR4"
FT Region 48..53 /note= "Tethered hexapeptide activating ligand
(AA50140)" Modified-site 56
FT /note= "N-glycosylated"
FT Domain 79..102 /note= "Transmembrane domain 1"
FT Region 103..110 /note= "Intracellular loop 1"
FT Domain 111..132 /note= "Transmembrane domain 2"
FT Region 133..150 /note= "Extracellular loop 1"
FT Domain 151..172 /note= "Transmembrane domain 3"
FT Region 173..191 /note= "Intracellular loop 2"
FT Domain 192..213

FT Region /note= "Transmembrane domain 4"
FT 214..239 /note= "Extracellular loop 2, determines specificity for activating peptide"
FT Region 228..230 /note= "These three residues are conserved among PAR1-4"
FT Domain 240..263 /note= "Transmembrane domain 5"
FT Region 264..283 /note= "Intracellular loop 3"
FT Domain 284..305 /note= "Transmembrane domain 6"
FT Region 306..316 /note= "Extracellular loop 3"
FT Domain 317..343 /note= "Transmembrane domain 7"
FT Region 344..385 /note= "Intracellular C-terminal region"
XX WO950415-A2.
XX PN 07-OCT-1999.
XX PD 31-MAR-1999; 99WO-US07100.
XX PF 01-APR-1998; 98US-0053866.
XX PR (ZYMO) ZYMOGENETICS INC.
XX PA (UNIW) UNIV WASHINGTON.
XX XX Xu W, Presnell SR, Yee DP, Foster DC;
PI WPI; 1999-633640/54.
XX DR N-PSDB; AA232747, AA232748.
XX PT Novel protease activated receptor 4, useful for screening for
FT (antagonists for promoting the proliferation and/or differentiation of
FT platelets and in mediating inflammatory events
XX Claim 8; Fig 1; 85pp; English.
XX CC This sequence represents a human protease-activated receptor, PAR4 (also
CC referred to as ZCHEMR2). The cDNA was identified from EST (expressed
CC sequence tag) sequences with homology to the three known protease-
CC activated receptors (PAR1, PAR2 and PAR3). Protease-activated receptors
CC (PARs) are a subfamily of G protein coupled receptors which are capable
CC of mediating cellular signalling in response to proteases (e.g.,
CC thrombin). They are characterised by a tethered peptide ligand at the
CC extracellular N-terminus that is generated by proteolysis. PAR4 is
CC activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates
CC a new N-terminus corresponding to the tethered ligand (a hexapeptide).
CC Agonists of PAR4 are useful for upregulating cellular or physiological
CC responses whereas antagonists are used to downregulate these
CC activities. The PAR4 protein is further useful for dissecting the
CC effects of thrombin or other activating proteases in the clotting
CC pathway from the effects of these proteases at the cellular level.
CC Agonists are specifically useful in promoting the proliferation
CC and/or differentiation of platelets, in mediating inflammatory events,
CC responses to vascular injury, chemotaxis or mitogenesis, and in
CC producing growth factors. Antagonists are useful as research reagents
CC for characterising sites of ligand-receptor interaction.
XX SQ Sequence 385 AA;
Query Match 24.2%; Score 521.5; DB 20; Length 385;
Best Local Similarity 33.8%; Pred. No. 1.9e-45;
Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;
QY 29 GRLLMLPLVLGFLSLSGGTQ-----TPSVYDESGSTGGDDSTPSILPAPRGYQGV 86
Db 3 GRLLMLPLVLGFLSLSGGTQ-----TPSVYDESGSTGGDDSTPSILPAPRGYQGV---- 53
QY 87 IKCPESASHLHVKNATMGYLTSSLSKLIPIAYLLVVFVGV PANAVTLMLFFRTRIC 146

Db 54 --CANDSDT--LELPDSSRALLGWPTRLVPALYGLVGLVGLPANGALWVLATQAPRLP 110
Qy 147 TTVFYTNLADELFVCTLPFKIAYHLNGNNWVGEVLCRAATVIFGNNMYCSILLACI 206
Db 111 STMLMNLATADLLALAPRIAYHLRGQRPFGAACLATAALYGHMYGVLAAV 170
Qy 207 SINRYLAIVHPFTYRGKHYALVTCGLVWATVFLYMLPFFILKOEYLVQPDITTC 266
Db 171 SLDRYLALVHLPLARALGRRLALGLCMAAWLMAALALPLTLQRTFLARSDRVLC 230
Qy 267 VHNTESSSPFQYFISLAFPGFLIPFVLIYCYAAIIRTNAYDHRMLWYVKASLL 326
Db 231 ALPLDAQSHWQ--PAFTCLALLGCFLLPALLMCLCYGATLHTLAASGRYGHALRL 289
Qy 327 VITICFAPSNIILIHANYNNYNTDGLYFIYLIACLGSLNSCLDPFLYFLMS 381
Db 290 ASAVAFVPSNLLLLHYSDPSAMGNLYGAYVPSLALSTLNSCVDPIFYVVS 344

RESULT 24
AA45036
ID AAY45036 standard; Protein; 385 AA.
AC AAY45036;
XX
XX 31-MAY-2000 (first entry)
XX Human protease activated receptor-4.
XX Human; PAR-4; protease activated receptor;
KW antisenase molecule; PAR antibody; cytostatic; therapeutic;
KW metastatic tumour cell; placental implantation; invasive cell.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 43..67
FT /note= "conserved unique region comprising protease
FT activated domain and hirudin binding domain.
FT Therapeutically useful antisenase molecules are designed
FT based on this region"
FT Region 228..230
FT /note= "conserved residues in second extracellular
FT loop used for designing therapeutically useful
FT antisenase molecules "
XX

PN WO200008150-A1.
XX
XX 17-FEB-2000.
XX
XX 05-FEB-1999; 99WO-IL00079.
XX
XX 07-AUG-1998; 98IL-0125698.
XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX Bar-Shavit R;
XX
XX WPI; 2000-205706/18.
DR N-PSDB; AA250775.
XX
XX Treating metastatic tumor cells useful for treating disorders involving
PT placenta implantation in a female comprises administration of an
PT antisenase molecule complementary to an RNA sequence of a protease
PT activated receptor protein -
XX
XX Example 3; Fig 11b; 46pp; English.
XX
XX The patent discloses a method to treat metastatic tumour cells using
CC an antisenase molecule comprising a polynucleotide complementary to an
CC RNA sequence of a protease activated receptor (PAR) protein, or an
CC antibody capable of binding to a PAR protein. The antisenase molecules and

CC antibodies of PAR protein are also used to treat disorders associated
CC with implantation of placenta. The present sequence is a human PAR-4
CC protein used for producing antisenase molecules for treating invasive
CC cells.
XX
SQ Sequence 385 AA;
Query Match 24.2%; Score 521.5; DB 21; Length 385;
Best Local Similarity 33.8%; Pred No 1.9e-45;
Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;
Qy 29 GLLLLLPFCQSGMENDTNLAKPTLPKTRFGAPPNSEFPF--FSALEGWGTATITVK 86
Db 3 GRLLMLPLVGLFSLGGTQ-----TPSVYDESGSTGGDDSTPSILPAPRGYPGVQV--- 53
Qy 87 IKCPESASHLVKVNATMGYLTSSLSSTKLIPALYLLVGVVYVYANAVTLMWLFRTSIC 146
Db 54 --CANDSDT--LELPDSSRALLGWPTRLVPALYGLVGLVGLPANGALWVLATQAPRLP 110
Qy 147 TTVFYTNLAIDFLFCVTLPFKIAYHLNGNNWVGEVLCRAATVIFGNNMYCSILLACI 206
Db 111 STMLMNLATADLLALAPRIAYHLRGQRPFGAACLATAALYGHMYGVLAAV 170
Qy 207 SINRYLAIVHPFTYRGKHYALVTCGLVWATVFLYMLPFFILKOEYLVQPDITTC 266
Db 171 SLDRYLALVHLPLARALGRRLALGLCMAAWLMAALALPLTLQRTFLARSDRVLC 230
Qy 267 VHNTESSSPFQYFISLAFPGFLIPFVLIYCYAAIIRTNAYDHRMLWYVKASLL 326
Db 231 ALPLDAQSHWQ--PAFTCLALLGCFLLPALLMCLCYGATLHTLAASGRYGHALRL 289
Qy 327 VITICFAPSNIILIHANYNNYNTDGLYFIYLIACLGSLNSCLDPFLYFLMS 381
Db 290 ASAVAFVPSNLLLLHYSDPSAMGNLYGAYVPSLALSTLNSCVDPIFYVVS 344

RESULT 25
AAB47623
ID AAB47623 standard; Protein; 385 AA.
XX
XX AAB47623;
XX 21-JAN-2002 (first entry)
XX Human PAR4.
XX Human; protease-activated receptor; PAR4; N-terminal; activated PAR4;
KW platelet aggregation; inhibition; tumour cell; proliferation.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 48..53
XX /label= N-terminal of activated PAR4
XX WO200158930-A1.
XX 16-AUG-2001.
XX
XX 06-FEB-2001; 2001WO-US03807.
XX
XX 09-FEB-2000; 2000US-0500646.
XX (ZYMO) ZYMOGENETICS INC.
XX Baidur N, West RR;
XX WPI; 2001-656678/75.
XX N-PSDB; AA43632.
XX Peptides comprising an amino acid sequence are capable of stimulating
PT protease-activated receptor 4 which are useful in diagnosis and therapy
PT e.g. inhibiting tumor cell proliferation and stimulating platelet

```
PT aggregation -
XX -PS Disclosure; Page 58-60; 84pp; English.
XX
XX This sequence shows human protease-activated receptor, PAR4. Peptides
CC derived from, or based on, the N-terminal of activated PAR4 (see
CC AAB47624-77), are capable of stimulating PAR4. These peptides may be
CC used to activate PAR4 at lower concentrations than wild type PAR4.
CC These peptides may be used to stimulate platelet aggregation, and for
CC inhibiting tumor cell proliferation.
XX
SQ Sequence 385 AA;
Query Match 24.2%; Score 521.5; DB 22; Length 385;
Best Local Similarity 33.8%; Pred. No. 1.9e-45;
Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;
QY 29 GLLLPTFCQSGMENDNNLAKPTLPKTRCAPNSPEEFP--PSALEGWTGATITVK 86
Db 3 GRULLPLVLGSLSGTQ-----TPSVYDESGTGGDDSTPSILPAPRGYQGQV---- 53
QY 87 IKCPESASHLVKNATMGVLTSSLSKLPATVLLVGVVPANAVTLMMLFFRTRSI 146
Db 54 --CANDSDT--LELPDSSRALLLGVWTRVLPALYGLVVLVGLPANGALWVLATQAPRLP 110
QY 147 TVVYNTLAIDFLFCVTLPFKIAYHLNGNWNVGEVLCRATTVI FYGNMYCSILLACI 206
Db 111 STMLLANLATADLLALPPIAYHLRQWRPFGEACRLATAALYGHMYGSVLLAAV 170
QY 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLMPLFFLKQEYVLVQDPITTC 266
Db 171 SLDRYLAIVHLPRARALGRRLALGLCMAALMAALALPLTLQROTFRLARSDRVLC 230
QY 267 VHTNCSSPFPOLYFISLAFFGLIPFVLIIYCYAAIIRTNAYDHRMLWYVKASILL 326
Db 231 ALPDAQASHWQ--PATCLLALGCFPLLANLLCYGATHTLAASGRRYGHARLFAVVL 289
QY 327 VFTICFAPSNIILIHANYNNYNTDGLFYIYIALCLGLNSCLDPPFLYFLMS 381
Db 290 ASAVAFVPSNLLLLHYSPPSAMGNLYGAVPSALSTLNSCVDPFYIYVVS 344
RESULT 26
AAV50138
ID AAY50138 standard; Protein; 385 AA.
XX
AC AAY50138;
XX
DT 31-JAN-2000 (first entry)
XX
DE Mutant human protease-activated receptor PAR4, R47A.
XX
KW Protease-activated receptor; PAR4; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;
KW platelet; proliferation; differentiation; mediation;
KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW growth factor; production; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PH Key Location/Qualifiers
FT Region 1..78 /note= "Extracellular N-terminal region"
FT Peptide 1..17 /note= "Signal peptide"
FT Cleavage-site 17..18 /note= "Cleaved by signal peptidase"
FT Protein 18..385 /note= "Mature non-activated human PAR4"
FT Region 48..53
/note= "Tethered hexapeptide activating ligand
/note= "Modified-site 56
/note= "N-glycosylated"
Domain
/note= "Transmembrane domain 1"
Region
/note= "Intracellular loop 1"
Domain
/note= "Transmembrane domain 2"
Region
/note= "Extracellular loop 1"
Domain
/note= "Transmembrane domain 3"
Region
/note= "Intracellular loop 2"
Domain
/note= "Transmembrane domain 4"
Region
/note= "Extracellular loop 2, determines specificity for activating peptide"
Region
/note= "These three residues are conserved among PAR1-4"
Domain
/note= "Transmembrane domain 5"
Region
/note= "Intracellular loop 3"
Domain
/note= "Transmembrane domain 6"
Region
/note= "Extracellular loop 3"
Domain
/note= "Transmembrane domain 7"
Region
/note= "Intracellular C-terminal region"
WO950415-A2.
07-OCT-1999.
31-MAR-1999; 99WO-US07100.
01-APR-1998; 98US-0053866.
(ZYMO ) ZYMOGENETICS INC.
(UNIW ) UNIV WASHINGTON.
Xu W, Presnell SR, Yee DP, Foster DC;
WPI; 1999-633640/54.
Novel protease activated receptor 4, useful for screening for
(ant)agonists for promoting the proliferation and/or differentiation of
platelets and in mediating inflammatory events -
Example 2; Page -; 85pp; English.
This sequence represents a mutant human protease-activated receptor PAR4
R47A, compared with PAR4 mutant R68A (AAV50139) in studies to evaluate
the importance of a protease cleavage site between residues 47 and 48.
This mutant PAR4 was unable to be cleaved by thrombin or trypsin, in
contrast to the wild-type (AAV50135) or the R68A mutant. This cleavage
activates PAR4, exposing a tethered hexapeptide ligand at the new
N-terminus. Protease-activated receptors (PARs) are a subfamily of G
protein coupled receptors which are capable of mediating cellular
signalling in response to proteases. Agonists of PAR4 are useful for
upregulating cellular or physiological responses whereas antagonists
are used to downregulate these activities. The PAR4 protein is
further useful for dissecting the effects of thrombin or other
activating proteases in the clotting pathway from the effects of these
proteases at the cellular level. Agonists are specifically useful in
promoting the proliferation and/or differentiation of platelets, in
mediating inflammatory events, responses to vascular injury, chemotaxis
or mitogenesis, and in producing growth factors. Antagonists are useful
```

CC as research reagents for characterising sites of ligand-receptor
CC interaction.
CC Note: This sequence is not shown in the specification, but is derived
CC from the sequence of the PA84 shown in figure 1.

Qx	Sequence	385 AA;
	Query Match	24.2%; Score 521; DB 20; Length 385;
	Best Local Similarity	36.9%; Pred. No. 2.e-45;
	Matches 108; Conservative	57; Mismatches 126; Indels 2; Gaps 2;
Qy	89 CPBESASHLHVKNATMGVLTSSLTKLIPALYLVFVGVGPANAVTLMLEFRTSRICTT	148
	: : : : : : : : : : : : : : : : : : : : : : : :	
Db	54 CANDSDT-LBEPDSSRALLLCGWPTRLVPALYGLVVLGVPANGIALWLVLTQAPRLPST	112
	: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	149 VFYTNLAIADELFCVTLPFKTIAYHLNGNNVFGVLCRAATTVIFGNNMCSILLACISI	208
	: : : : : : : : : : : : : : : : : : : : : : : :	
Db	113 MLLMNLTAIDLALLALPPRIAYHLRQORWPFGEAACRLTAALYGHMVGSVLLAAVSL	172
	: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	209 NRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFTLKQEYLVLPQDITCTCHDVH	268
	: : : : : : : : : : : : : : : : : : : : : : : :	
Db	173 DRYLALVHPLEARALRGRRLLGLGMAAWLMAAALALPLTLQRTFLARSDRVLCHDAL	232
	: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	269 NTCBSSPFFOLYYPISLAFFGFLIPFVLIIYCYAAIIRTLNAYDHRWLWVYKVALIIVI	328
	: : : : : : : : : : : : : : : : : : : : : : : :	
Db	233 PLDAQASHMQ-PAFTCLALLGCFPLLLAMLLCYGATLHTLAASGRRYGHALRLTAVVLAS	291
	: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	329 FTTCFAPSNIILITHANYNNYNTDGLYFIVYLIALCGLSNLSCLDDPFYLFYFMS	381
	: : : : : : : : : : : : : : : : : : : : : : : :	
Db	292 AVAFFVSNILLLHYDPDPSAAGNLYGAVVPSLALSTLNSCYDVPFYFYFVS	344
	: : : : : : : : : : : : : : : : : : : : : : : :	

RESULT 27
ABG35300
ID ABG35300 standard; Protein; 425 AA.
XX
XX AC
XX AC
XX ABG35300;
XX
XX
DT 15-JUL-2002 (first entry)

XX	
KW	Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW	inflammatory disease; cell proliferative disease.
XX	
OS	Homo sapiens.

XX
PN JP2002010784-A.
XX
XX
PD 15-JAN-2002.
XX
XX 29-JUN-2000; 2000JP-0196514.
XX
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
XX (TEIJ) TEIJIN LTD.
XX
XX
XX WPI; 2002-321520/36.
DR N-PSDB; ABK70889.
DR

PT An inhibitor of cell growth mediated by thrombin used to treat inflammatory and cell proliferative diseases -

PS Disclosure; Page 26-27; 44pp; Japanese.

The invention relates to a polypeptide or a compound which can inhibit cell growth caused by thrombin. The polypeptide/compound combines to a specific region of the structure of PAR1 type human thrombin receptor participating to cell growth. Preferably, the compound contains the 52nd to the 56th amino acid sequences at the amino end side of PAR1 type human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Irp-(X5) X4, X5 = optional amino acid or peptide sequence). Also included are a modified PAR1 type thrombin receptor gene or its fragment used for obtaining the above

CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
CC version).

XX	SQ	Sequence	425 AA;
		Query Match	24.1%; Score 519; DB 23; Length 425;
		Best Local Similarity	32.6%; Pred. No. 3.9e-45;
		Matches	126; Conservative 76; Mismatches 154; Indels 30; Gaps 11;
Qy	14	GTQVIKWKALIPAAAGULLLLPFCOSGMENDTNNLAKPTLPKIKTRGAPPN-SPREPPF	72
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	2	GPRLLLVAAACFLSCGLPLSARTRARRPESKATN-----ATLDRSELNPNNDKYEPFWE	57
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	73	SALEGMTGAT-----ITVKIKCPBESASHLVKNATMGYLTSSTLKLIPAIYLLLVFVGV	128
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	58	DEEKNESGLTEYRLVSINKSSPLQKLPALFISEDASGYLTSSWTLTPVPSVYTGVFVWSL	117
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	129	PAN--AVTLWMLPFRTRSTCTTVFYNNLAIADFLCVTLPPFKIAYHLNGNNWVFGVLCR	186
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	118	PLNMAIVAVFILKMKYKK-PAVVYMLHLATADVLFSVLPFKISYFSGSDWQFSELCL	176
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	187	ATTVIFGNNYCSILLACISINRYLAIVHP-----FYRGLPKHYALVTCGLVNAWTVFLY	243
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	177	FVTAAYFCNMYASILLMTVISIDRFVAVVYPMQSLSWRTLGRASF---TCLAIWALAIAG	233
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	244	MLPFFILKQBYLVQDDITTCCHDVHNTCESSSPQIYYFISLAFQGLPFPVLLIYCYAA	303
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	234	WPPELLKEQTIQVGNLITTCCHDVLENETLEG--YYAYYFSAFSAVFVFLIISTVCYVS	292
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	304	IITLTN-----AYDHRMLWYWKASLLILVITFCIFAPSNIILIIHHANY--YNNNTDGL	355
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	293	IICLSSAVANKSKSRALF---LSAAVFCPIICFGPTNVLIIAHYFSLSHSTTEAA	349
		: : : : : : : : : : : : : : : : : : : : : : : :	
Qy	356	YFYTLIALCGLSUNSCLDPLPYFLMS	381
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	350	YFAYLLCVCVSSISCDDILYIYYAS	375
		: : : : : : : : : : : : : : : : : : : : : : : :	

RESULT 28	
AAR60698	
ID	AAR60698 standard; Protein; 425 AA.
XX	
XX	
AC	AAR60698;
XX	
XX	
DT	25-JUN-1995 (first entry)
XX	
DE	Fragment of the human thrombin receptor.

PI Betlach MC, Turner GJ;

XX
DR WPI; 1994-317010/39.

DR N-PSDB; AAQ/3390.
XX

PT Expression of heterologous proteins in halo-bacteria - using
PT regulatory and stop sequences from halo-bacteria, pref. the
PT bacterio-rhodopsin gene.

```
XX Disclosure; Fig 14; 118pp; English.
XX
XX The sequence is that of the human thrombin receptor fragment.
CC This is used to exemplify a new expression vector for producing
CC heterologous polypeptides in a halobacterial host.
CC See also AAR60891-9.
XX
SQ Sequence 425 AA;
Query Match 23.9%; Score 516.5; DB 15; Length 425;
Best Local Similarity 32.8%; Pred. No. 7.2e-45; Indels 33; Gaps 11;
Matches 123; Conservative 77; Mismatches 144;
XX
XX 30 LLLLPFTFCQ-----SGMENDTNNAKPTLPKTRGAPPN-SPEEPFSALEGTGA 81
Db 1 MLELLPTAVEGVSAQIQALDDDDVDATLDRSFLLRNPNDKYEPFWEDEKNESGL 60
QY 82 T-----ITVKIKPEESASHLVKNATMGVLTSSLSKLPAIYLLVFGVGPAN--AVTL 135
Db 61 TEYRLVSINKSSPQKLPAPFISEDASGYLTSSWLTLPVPSVYTGVPVSLPNIIMAVV 120
QY 136 WMLPFFTRISCTTVEFTNLAIADELFCVTLPPKIAVHLNGNNVFGVLCRATTIVFGN 195
Db 121 FILKMKVKK-PAVYVMLHLATADVLSVLPFKISYFGSDMQFGSELRCFVTAAPYCN 179
QY 196 MYCSILLACISINRYLAIVHP---FYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQ 252
Db 180 MYASILLMTVISIDFLAVVPMQSLSWRTLGRASF---TCLAIWALAIAGVVPVLVKEQ 236
QY 253 EYVLVQPDITTDCHDVNTCESSPPFQYFISLAFFGLIPFVLIYCYAAIIRTLN--- 309
Db 237 TIQVPLNITTDCHDVNLTLLEG-YYAYYFSAFSAVFFVPLIISTVCVYSIIRCLSSA 295
QY 310 ----AYDHRWLWVVKASLLILVFTICFAPSNIILIIHHANY-YNNTDGLYFIYLIAC 364
Db 296 VANRSKKSRALF---LSAAVFCIFICFGPTNVLIIAHYSFUSHTSTTEAAYFAYLLVC 352
QY 365 LGSLSNCLDPFLYFLMS 381
Db 353 VSSISSCIDPLIYYVAS 369
RESULT 29
AAW16314
XX AAW16314 standard; Protein; 892 AA.
AC AAW16314;
XX
XX 16-AUG-1997 (first entry)
DE Human thrombin receptor-yeast G-alpha protein fusion.
XX
XX G-protein coupled receptor; agonist; antagonist; assay;
KW G-alpha protein; Gpalp; GPAL gene; thrombin receptor.
XX
XX Chimaeric Homo sapiens;
OS Chimaeric Saccharomyces cerevisiae.
XX
XX Key Location/Qualifiers
FT Peptide 1..13
FT /label= Sig_peptide
FT /note= "alpha-factor (STE2) signal peptide"
FT Protein 14..892
FT /label= Thr-GPAL_fusion
FT Region 14..417
FT /label= Thr
FT /note= "thrombin receptor amino acids 22-425"
FT Peptide 418..420
FT /label= Linker
FT Region 421..892
FT /label= G-alpha_protein
XX
```

```
PN WO9711159-A1.
XX
XX 27-MAR-1997.
XX
XX 20-SEP-1996; 96WO-US15203.
XX
XX 20-SEP-1995; 95US-0004023.
XX
XX (HEAR-) HEARTLAND BIOTECHNOLOGIES LLC.
XX
XX Boulton TG, Das PO, Mandell RB, McMullen TW;
XX WPI; 1997-202868/18.
DR N-PSDB; AAT62461.
XX
XX New yeast constructs for use in screening assays - which express a
PT fusion comprising a mammalian or fungal transmembrane receptor and
PT a yeast G-alpha protein
XX
XX Example 2; Fig 12A-G; 109pp; English.
XX
XX A fusion protein (AAW16314) comprises the human thrombin receptor
CC (amino acids 22 to C-terminal threonine) covalently linked to
CC yeast G-alpha protein Gpalp (encoded by the GPAL gene). It is
CC expressed by a Thr-GPAL gene fusion (AAT62461) in vector pRMBT16.
CC Thrombin-dependent activation of the yeast mating pathway by the
CC Thr-GPAL fusion protein was observed in ste2 gpal- yeast cells.
CC The polypeptide is an example of novel fusion proteins between
CC eukaryotic heterotrimeric G-protein coupled receptors and yeast
CC G-alpha proteins that can be coupled to the pheromone-induced
CC signal transduction pathway of yeast and used in screening assays
CC to identify agonists or antagonists of the receptor.
XX
XX Sequence 892 AA;
Query Match 23.7%; Score 512; DB 18; Length 892;
Best Local Similarity 34.0%; Pred. No. 5.5e-44;
Matches 118; Conservative 71; Mismatches 132; Indels 26; Gaps 10;
QY 53 TLPIKTRGAPPN-SPEEPFSALEGTGAT---ITVKIKPEESASHLVKNATMGYL 107
Db 29 TLDRSFLLRNPNDKYEPFWEDEKNESGLTEYRLVSINKSSPQKLPAPFISEDASGYL 88
QY 108 TSSLSTKLIPAIYLLVFGVGPAN--AVTLWMLFFTRISCTTVEFTNLAIADELFCVTL 165
Db 89 TSSWLTLPVPSVYTGVPVSLPNIIMAVIFILKMKVKK-PAVYVMLHLATADVLSVVL 147
QY 166 PKIAYHLNGNNVFGVLCRATTIVFYGNMYCSILLACISINRYLAIVHP---FTYRG 222
Db 148 PFKISYFSGSDMQFGSELRCFVTAAFYCNMYASILLMTVISIDRFLAVVPMQSLSWRT 207
QY 223 LPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTDCHDVNTCESSPPFOLY 282
Db 208 LGRASF---TCLAIWALAIAGVVPVLVKEQTITQVPLNITTDCHDVNLTLLEG-Y 263
QY 283 ISLAFFGLIPFVLIYCYAAIIRTLN-----AYDHRWLWVVKASLLILVFTICFAP 335
Db 264 SAFAVFFVPLIISTVCVYSIIRCLSSAVANRSKKSRALF---LSAAVFCIFICFGP 320
QY 336 SNIIILIIHHANY-YNNTDGLYFIYLIACLGSLNSCLDPFLYFLMS 381
Db 321 TNVLLIIAHYSFLSHTSTTEAAYFAYLLVCVSSISSCIDPLIYYVAS 367
RESULT 30
ABG35298
ID ABG35298 standard; Protein; 402 AA.
XX
XX ABG35298;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human PAR1 type thrombin receptor delta 1-49TR.
```

XX Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX Homo sapiens.

OS

XX JP2002010784-A.

XX 15-JAN-2002.

XX 29-JUN-2000; 2000JP-0196514.

XX 29-JUN-2000; 2000JP-0196514.

XX (TEIJ) TEIJIN LTD.

XX WPI; 2002-321520/36.

DR N-PSDB; ABK70887.

XX An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -

XX Disclosure; Page 22-24; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
CC version).

XX Sequence 402 AA;

Query Match 23.5%; Score 507.5; DB 23; Length 402;
Best Local Similarity 33.0%; Pred. No. 5.7e-44;
Matches 121; Conservative 71; Mismatches 138; Indels 37; Gaps 11;

QY 32 LLLPTFCQSGMENDNNLAKPTLPIKTRGAPNSPEEPFSALEGWGTGAT-----ITVKI 87

DB 6 LLLVAACFS-----LCGLPLSART-----RADKYEPPFWEDEKNEGSLTEYRLVSINK 53

QY 88 KCPEESASHLRVKNATNGYLTSLTKLIPAIYLLVFGVGVAN--AVTLWMLFFFRTRSI 145

DB 54 SSPLQKLPAPFISEDASGYLTSSWTLTFVPSVTVGVVSLPLNINAIIVFVILKMKVKK- 112

QY 146 CTTVFYNNLAIDFLFCVTLPPKIAHYLNGNNWVGEVLCRATTVIFYGNMYSIILLAC 205

DB 113 PAVYMLHLATADVLVSVLPFKISYFSGSDQFSGELCRFVTAAYFCNMYASILLMTV 172

QY 206 ISINRYLAIVHP---FTYRGLPKHTALVTCGLVWATVFLYMLPFFILKQYVLPQDIT 262

DB 173 ISIDRFVAVVYPMQSLWSRITLGRASF---TCLAIWALAIAGVVPVLLKQETIQVPGNLT 229

QY 263 TCHDVHNTCESSPPFQLYYISIAFFGLIPFVLIICYAAIIRTLN-----AYDHRW 315

DB 230 TCHDVNETLEGG-YAYAYESAFSAVFFVPLIISTVCYVSIIRCLSSSAVANRKSRA 288

QY 316 LWYKASLLILVFTICFAPSNIILIIHANY-YNNNTDGLFYIYIALCLGSLNSCLDP 374

DB 289 LF---LSAAVFCIFIIICFGPTNLLIAHYSFLSHTSTTEAAYFAYLLVCVSSISCCIDP 345

QY 375 FLYFLMS 381

DB 346 LIYYAS 352

RESULT 31

ABG35299

ID ABG35299 standard; Protein; 371 AA.

XX AC ABG35299;

XX 15-JUL-2002 (first entry)

XX Human PAR1 type thrombin receptor delta 1-80.

XX Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.

XX Homo sapiens.

XX JP2002010784-A.

XX 15-JAN-2002.

XX 29-JUN-2000; 2000JP-0196514.

XX 29-JUN-2000; 2000JP-0196514.

XX (TEIJ) TEIJIN LTD.

XX WPI; 2002-321520/36.

XX N-PSDB; ABK70888.

XX An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -

XX Disclosure; Page 24-25; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
CC version).

XX Sequence 371 AA;

Query Match 23.2%; Score 499.5; DB 23; Length 371;
Best Local Similarity 36.9%; Pred. No. 3.5e-43;
Matches 107; Conservative 61; Mismatches 101; Indels 21; Gaps 8;

QY 105 GYLTSLSLTKLIPAIYLLVFGVGVAN--AVTLWMLFFFRTRISCTTVFVNNLAIDFLFC 162

DB 40 GYLTSLSLTKLIPAIYLLVFGVGVAN--AVTLWMLFFFRTRISCTTVFVNNLAIDFLFC 98

QY 163 VTLPPKIAHYLNGNNWVGEVLCRATTVIFYGNMYSIILLACISINRYLAIVHP---FT 219

DB 99 SVLPFKISYFSGSDQFSGELCRFVTAAYFCNMYASILLMTVISIDRFVAVVYPMQSL 158

QY 220 YRGLPKHTALVTCGLVWATVFLYMLPFFILKQYVLPQDITTCCHDVHNTCESSPPFQL 279

DB 159 WRTLGRASF---TCLAIWALAIAGVVPVLLKQETIQVPGNLTTCCHDVHNTLEGG-Y 214

QY 280 YFISIAFFGLIPFVLIICYAAIIRTLN-----AYDHRWLVWYKASLLILVFTIC 332

DB 215 YFISIAFFGLIPFVLIICYAAIIRTLN-----AYDHRWLVWYKASLLILVFTIC 271

QY 333 FAPSNIILIIHANY-YNNNTDGLFYIYIALCLGSLNSCLDPFLYFLMS 381

DB 272 FGPTNLLIAHYSFLSHTSTTEAAYFAYLLVCVSSISCCIDPLIYYAS 321

QY	232	TCGLVWATVFLYMLPFILKQBYLVLPDITTCCHDVHNTC-----ESSSPFQLYYF	28
Db	171	LCIVAVLSAATLALPLTLHRQNFLLAP-IACC--VMRCPLMSTRPTGERPSSAWLSWA	227
QY	283	ISLAFFGFLIPFVLIYCYAAIIRTLNAYDHWLWVVKASLLILVITICFAPSNIILII	342
Db	228	ASL-----PLAMGLCYGTTIRALAANGQRYSHALRALTALVLFSAVASFTPSNVLLV	280
QY	343	HHANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMS	381
Db	281	HYSNPSPSEAGNLYGAYVPSLALSTLNSCVDPIFYIYVS	319
RESULT 33			
AAW71304			
ID	AAW71304	standard; Protein; 359 AA.	
XX	XX	AAW71304;	
AC	XX	AAW71304;	
XX	XX	02-NOV-2000 (first entry)	
DT	XX	Human orphan G protein-coupled receptor hCHN4.	
XX	XX	Human; orphan G protein-coupled receptor; GPCR; hCHN4; drug screening;	
KW	XX	transmembrane receptor; expressed sequence tag; EST; signal cascade.	
KW	XX	Homo sapiens.	
OS	XX	WO200031258-A2.	
XX	XX	02-JUN-2000.	
PN	XX	13-OCT-1999; 99WO-US23687.	
PD	XX	20-NOV-1998; 99US-0109213.	
XX	XX	16-FEB-1999; 99US-0120416.	
PR	XX	26-FEB-1999; 99US-0121852.	
PR	XX	12-MAR-1999; 99US-0123946.	
PR	XX	12-MAR-1999; 99US-0123949.	
PR	XX	28-MAY-1999; 99US-0136436.	
PR	XX	28-MAY-1999; 99US-0136437.	
PR	XX	28-MAY-1999; 99US-0136439.	
PR	XX	28-MAY-1999; 99US-0136567.	
PR	XX	28-MAY-1999; 99US-0137127.	
PR	XX	28-JUN-1999; 99US-0137131.	
PR	XX	29-JUN-1999; 99US-0141448.	
PR	XX	29-SEP-1999; 99US-0156555.	
PR	XX	29-SEP-1999; 99US-0156633.	
PR	XX	29-SEP-1999; 99US-0156634.	
PR	XX	29-SEP-1999; 99US-0156653.	
PR	XX	01-OCT-1999; 99US-0157280.	
PR	XX	01-OCT-1999; 99US-0157281.	
PR	XX	01-OCT-1999; 99US-0157282.	
PR	XX	01-OCT-1999; 99US-0157293.	
PR	XX	01-OCT-1999; 99US-0157294.	
PR	XX	12-OCT-1999; 99US-0416760.	
PR	XX	12-OCT-1999; 99US-0417044.	
XX	XX	(AREN-) ARENA PHARM INC.	
XX	XX	Chen R, Dang HT, Liaw CW, Lin I;	
PI	XX	WPI; 2000-400068/34.	
DR	XX	N-PSDB; AAD01131.	
XX	XX	Novel human orphan G protein-coupled receptors and the encoding cDNAs	
PT	XX	for use in the identification of G protein-coupled receptor agonists -	
XX	XX	Claim 54; Page 77-78; 102pp; English.	
XX	XX	The present amino acid sequence is the hCHN4, an endogenous human	
CC	XX	orphan G protein-coupled receptor (GPCR). The hCHN4 cDNA was identified	

XX Human; G-protein coupled receptor; HG52; immune system; thrombin;
KW fibrinogen; fibrin; clotting factor; procoagulant; platelet activation;
KW chemotaxis; mitogenesis.
XX
OS Homo sapiens.
XX
XX W0200020438-A1.
XX
XX 13-APR-2000.
XX
XX 29-SEP-1999; 99WO-US22634.
XX
XX 02-OCT-1998; 98US-0102958.
XX
XX (MERI) MERCK & CO INC.
XX
XX Liu Q, McDonald TP, Wang R;
XX
XX WPI; 2000-317696/27.
XX
XX N-PSDB; AAA14828.
XX
XX New recombinant DNA encoding a G-protein coupled receptor designated
PT HG52 is useful to find modulators of thrombin effects and shows
PT homology to thrombin receptors -
XX
XX Claim 6; Fig 2; 39pp; English.
XX
XX The present sequence represents a human G-protein coupled receptor
CC designated HG52. HG52 is a member of the rhodopsin family. The HG52
CC RNA is widely expressed in humans as a transcript of about 4.5. kb,
CC especially in cells of the immune system. The HG52 DNA can be used
CC in chromosomal mapping studies, and to identify individuals carrying
CC a disease-carrying gene. Agonists and antagonists of HG52 will be
CC useful as modulators of the effects of thrombin, including conversion
CC of fibrinogen to fibrin in plasma, activation of clotting factors V,
CC VIII, XIII and protein C, the procoagulant function of platelets and
CC endothelial cells, stimulation of platelet activation, chemotaxis of
CC monocytes and lymphocytes, and mitogenesis of lymphocytes and
CC mesenchymal cells such as vascular smooth muscle cells, fibroblasts
CC and epithelial cells.
XX
XX Sequence 359 AA;
XX
Query Match 21.3%; Score 460; DB 21; Length 359;
Best Local Similarity 34.4%; Pred. No. 4.2e-39;
Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;
QY 101 NATMGYLTSSLTKLIPAIYLLVFGVGPANAVTLWMLFPR--TRSICTTVFVTNLAID 158
Db 11 NATLQMLRNPAIAVALPVYSLVAAVSIPIGNLFSVLVLCRRMGPRS-PSVIFMINLSVTD 69
QY 159 FLFCVTLPEFKIAYHLNGNNVGEVLCRATVIFYGNMYCSILLACISINRYLAIVHPF 218
Db 70 LMLASVLPFOIYHCNRRHHVFGVLLCNVTVAFYANNYSIIITMTCSIVERFLGVLYPL 129
QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEVYLVQPDITTCDDVHN-TCESSSPF 277
Db 130 SSRKRRRRVAVACAGTWWLLLTALSPLARTDLTPVHALGHTICPDVUKWTMLPSVAM 189
QY 278 QLYYFISLAFPGFELIPVLLIYCYAA----IIRLNAYD-HRWLWYVKASLLIIVFTIC 332
Db 190 WAVFLFTIFILLFPIPVITVACVATILKLLRTEEAHGRRRAVGLAAVLLAFVTC 249
QY 333 FAPSNIIILIH-HANYNNYNTDGLYFIYILALCLGSLNSCLDPLFLMSK 382
Db 250 FAPNNFVLLAHIVSRLPYFGKS--YYHYVKLTCLCLNCLNCLDPPVYFYFASR 298
RESULT 36
AAY69485
ID AAY69485 standard; Protein; 359 AA.
XX

AC AAY69485;
XX
XX 03-JUL-2000 (first entry)
XX
XX Amino acid sequence of a human 14400 receptor polypeptide.
XX
XX Human; 14400 receptor; G-protein coupled receptor; signalling pathway;
KW CD8 T cell; CD4 T cell; CD34+ bone marrow cell; thrombocytopenia;
KW inflammation; spleen disorder; splenomegaly; lung disorder;
KW adult respiratory distress syndrome; colon disorder;
KW bacterial enterocolitis; liver disorder; hepatic injury;
KW platelet number; precursor T-cell neoplasm.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 1..23 /note= "extracellular domain"
XX Domain 24..296 /note= "transmembrane domain"
XX Domain 120..122 /note= "GPCR signal transduction signature"
XX FT Misc-difference 278 /note= "Thr encoded by AGC"
XX FT Domain 297..359 /note= "intracellular domain"
XX
XX W0200011170-A1.
XX
XX 02-MAR-2000.
XX
XX 20-AUG-1999; 99WO-US19112.
XX
XX 20-AUG-1998; 98US-0137063.
XX 20-AUG-1999; 99US-0378100.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Glucksmann MA, Weich NS;
XX
XX WPI; 2000-256378/22.
XX N-PSDB; AA299588.
XX
XX G-protein coupled receptor, 14400 receptor, for treating CD4, CD8 or
XX CD34+ related disorders -
XX
XX Claim 1; Fig 1A-B; 92pp; English.
XX
XX The present sequence represents a human 14400 receptor polypeptide. The
XX receptor is a G-protein coupled receptor that participates in signalling
XX pathways. The polypeptide can be used to identify an agent to modulate
XX the polypeptides activity in patients having a disorder involving CD8 or
XX CD4 T cells or CD34+ bone marrow cells, especially where the disorder
XX involves thrombocytopenia or inflammation. The modulators can
XX be used to treat disorders involving the spleen (e.g. splenomegaly),
XX lung (e.g. adult respiratory distress syndrome), colon (e.g. bacterial
XX enterocolitis), liver (e.g. hepatic injury), the uterus and
XX endometrium, brain, T-cells, skin, heart, blood vessels, red cells,
XX thymus, B-cells, kidney, breast, testis and epididymis, prostate,
XX thyroid, skeletal muscle, pancreas, small intestine, reduced platelet
XX number, precursor T-cell neoplasms, CD3, CD4, and CD8 T lymphocytes.
XX
XX Sequence 359 AA;
XX
Query Match 21.3%; Score 460; DB 21; Length 359;
Best Local Similarity 34.4%; Pred. No. 4.2e-39;
Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;
QY 101 NATMGYLTSSLTKLIPAIYLLVFGVGPANAVTLWMLFPR--TRSICTTVFVTNLAID 158
Db 11 NATLQMLRNPAIAVALPVYSLVAAVSIPIGNLFSVLVLCRRMGPRS-PSVIFMINLSVTD 69
QY 159 FLFCVTLPEFKIAYHLNGNNVGEVLCRATVIFYGNMYCSILLACISINRYLAIVHPF 218

Db 70 LMLASVLPQIYHCHNRHHVFGVLLCNVTVAFYANMYSSILMTWCISVERFLGVLYPL 129
 QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVLYVQPDITTCDDVHN-TCSSSSPF 277
 Db 130 SSKRWRRRYAAVACAGTWLLLTALSPARTDLYTPVHALGIITCFDLKWTMLPSVAM 189
 QY 278 QLYYFSLAPFGFLPELVLIYCAA----IIRTLNAYD-HRWLVWVKASLLILVFTIC 332
 Db 190 WAVFLTFIFLLFPFVITVACYTATILKLTTEAHGREQSAAGVLAUVLLAFVTC 249
 QY 333 FAPSNIIILIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
 Db 250 FAPNPFVLLAHIVSRIFYGKS--YYHVYKLTCLCLNCLNCLDPFVYFASR 298

RESULT 37
 AAG78530
 ID AAG78530 standard; Protein; 359 AA.
 XX
 AC AAG78530;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE G-protein coupled receptor (designated Paul) amino acid sequence.
 XX
 KW G-protein coupled receptor; paul; 7TM receptor; GPC receptor;
 KW human; antibacterial; virucide; fungicide; protozoacide;
 KW anti-Human immunodeficiency virus; analgesic; cytostatic; antidiabetic;
 KW anorectic; antiaesthetic; antiparkinsonian; cardiant; hypertensive;
 KW hypotensive; osteopathic; antianigmal; antimanic; cerebroprotective;
 KW antiulcer; antiallergic; antidepressant; antimigraine; antiemetic;
 KW tranquiliser; neuroleptic; neuroprotective; nootropic; anticonvulsant;
 KW Human immunodeficiency virus type 1; HIV-2; HIV-1; pain; cancer;
 KW diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease;
 KW acute heart failure; hypertension; myocardial infarction; stroke; ulcer;
 KW osteoporosis; angina pectoris; myocardi retention;
 KW allergy; benign prostatic hypertrophy; migraine; vomiting; anxiety;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.
 XX
 OS Homo sapiens.
 XX
 PN US2001029032-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 18-DEC-2000; 2000US-0739151.
 XX
 PR 09-OCT-1998; 98US-103789P.
 PR 06-OCT-1999; 99US-0413534.
 XX
 PA (ZHU)/ ZHU Y.
 PA (LIXX)/ LI X.
 PA (VAVT)/ VAWTER L.
 XX
 PI Zhu Y, Li X, Vawter L;
 XX
 DR WPI; 2001-647985/74.
 DR N-PSDS; AAI64231.
 XX
 PT New G-protein coupled receptor polypeptide, referred as Paul and
 PT encoding polynucleotide, useful for diagnosing and treating cancers,
 PT infections, neurological disorders, diabetes, asthma and identifying
 PT modulators -
 XX
 PS Claim 1c; Page 13-14; 15pp; English.
 XX
 CC The invention relates to an isolated polypeptide, a member of G-protein
 CC coupled receptor family of polypeptides, comprising a fully defined
 CC sequence of 359 amino acids, its 95% identical sequence, or a polypeptide
 CC encoded by a polynucleotide comprising a fully defined sequence of 1080
 CC base pairs defined in the specification, or a fragment or variant of it.

CC The activity of the protein of the invention may be described as
 CC antibacterial, virucide, fungicide, protozoacide, anti-Human
 CC immunodeficiency virus, analgesic, cytostatic, antidiabetic, anorectic,
 CC antiaesthetic, antiparkinsonian, cardiant, hypertensive, hypotensive,
 CC osteopathic, antianigmal, antimanic, cerebroprotective, antiulcer,
 CC antiallergic, antidepressant, antimigraine, antiemetic, tranquiliser,
 CC neuroleptic, neuroprotective, nootropic and anticonvulsant. Polypeptides
 CC and polynucleotides of the invention are useful in diagnosis and in
 CC identifying compounds such as agonists and antagonists which are useful
 CC in therapy. They are also useful for treating diseases, including
 CC infections such as bacterial, fungal, protozoan and viral infections,
 CC particularly infections caused by Human immunodeficiency virus type 1 or
 CC 2 (HIV-1 or HIV-2), pain, cancers, diabetes, obesity, anorexia, bulimia,
 CC asthma, Parkinson's disease, acute heart failure, hypotension,
 CC hypertension, urinary retention, osteoporosis, angina pectoris,
 CC myocardial infarction, stroke, ulcers, allergies, benign prostatic
 CC hypertrophy, migraine, vomiting, psychotic and neurological disorders,
 CC including anxiety, schizophrenia, manic depression, depression, delirium,
 CC dementia, and severe mental retardation, dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome. Polypeptides and
 CC polynucleotides of the invention may also be useful as diagnostic
 CC reagents, for example in the detection of mutations in the associated
 CC gene, chromosome localisation studies and expression pattern
 CC determination. The current sequence represents a G-protein coupled
 CC receptor (designated "Paul" in the specification) amino acid sequence.
 XX
 SQ Sequence 359 AA;
 Query Match 21.3%; Score 460; DB 22; Length 359;
 Best Local Similarity 34.4%; Pred. No. 4.2e-39;
 Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;
 QY 101 NATMGVLTSSSLTKLIPALYLLAVFVGVUPANAVTLMLEFR--TBSICTTVFVTMLAIAD 158
 Db 11 NATLOLRNPAJAVALPVVYSLVAAVSPGNLFLWLCRRMGPRS-PSVIFMINLSVTD 69
 QY 159 FLFCVTLPPKIAHYHNGNNWFEVLCRATTVFYGNMYCSILLACISINRYLAIVHPF 218
 Db 70 LMASLVLPQIYHCHNRHHVFGVLLCNVTVAFYANMYSSILMTWCISVERFLGVLYPL 129
 QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVLYVQPDITTCDDVHN-TCSSSSPF 277
 Db 130 SSKRWRRRYAAVACAGTWLLLTALSPARTDLYTPVHALGIITCFDLKWTMLPSVAM 189
 QY 278 QLYYFSLAPFGFLPELVLIYCAA----IIRTLNAYD-HRWLVWVKASLLILVFTIC 332
 Db 190 WAVFLTFIFLLFPFVITVACYTATILKLTTEAHGREQSAAGVLAUVLLAFVTC 249
 QY 333 FAPSNIIILIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPFLYFLMSK 382
 Db 250 FAPNPFVLLAHIVSRIFYGKS--YYHVYKLTCLCLNCLNCLDPFVYFASR 298

RESULT 38
 AAG80966
 ID AAG80966 standard; Protein; 359 AA.
 XX
 AC AAG80966;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Human, nGPR5 #2.
 XX
 KW G protein-coupled receptor; nGPR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective.

OS Homo sapiens.
 XX WO200136473-A2.
 XX 25-MAY-2001.
 XX 16-NOV-2000; 2000WO-US31581.
 XX 16-NOV-1999; 99US-0165838.
 XX 17-NOV-1999; 99US-0166071.
 XX 19-NOV-1999; 99US-0166678.
 XX 28-DEC-1999; 99US-0173396.
 XX 22-FEB-2000; 2000US-0184129.
 XX 28-FEB-2000; 2000US-0185421.
 XX 28-FEB-2000; 2000US-0185554.
 XX 02-MAR-2000; 2000US-0186530.
 XX 03-MAR-2000; 2000US-0186811.
 XX 09-MAR-2000; 2000US-0188114.
 XX 17-MAR-2000; 2000US-0190310.
 XX 21-MAR-2000; 2000US-0190800.
 XX 20-APR-2000; 2000US-0198568.
 XX 02-MAY-2000; 2000US-0201190.
 XX 08-MAY-2000; 2000US-0203111.
 XX 25-MAY-2000; 2000US-0207094.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J,
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI; 2001-389826/41.
 XX N-PSDB; AAHS1006.
 XX New G protein-coupled receptor (ngPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX Claim 37; Page 88; 261pp; English.
 XX The present invention relates to novel G protein-coupled receptors
 CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is one such G protein-coupled receptor. GPCR's are also known as
 CC seven transmembrane receptors and function in signal transduction. The
 CC ngPCRx coding sequences are useful for screening a human to diagnose a
 CC disorder affecting the brain or a genetic predisposition, specifically
 CC schizophrenia. ngPCRx are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of ngPCRx in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of ngPCRx activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. inflammatory bowel disease.
 XX Sequence 359 AA;
 XX
 Query Match 21.3%; Score 460; DB 22; Length 359;
 Best Local Similarity 34.4%; Pred. No. 4.2e-39;
 Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;
 QY 101 NATGTYLTSLSKLIPAIYLLVFVGVDPANAVTLWMLFRR--TRSICTTFTVFNIAIAD 158
 DB 11 NATLQMRPAIAVALPWTSLVAASVPGNLSLWLCRRMGPRS-PSVIFMNLSTVD 69
 QY 159 FLECVTLPFKIAHLNGNNVFGVLCRATTVIFGNNVCSILLACISINRVAIVHPF 218
 DB 70 LMLASLVPFQIYYHCNRHHVFGVLLCNVTVTAFYANNVSSILMTCTISVERFLGVLYPL 129

QY 219 TYRGLPKHTYALVTGCLVWATVFLVWLPFFILKQEYVLVQPDITTTCHDVHN--TCSESSSPF 277
 DB 130 SSKRWRRRRYAAACAGTMLLLLTALSPILARTDLYPVHALGIITCFDLKWTMLPSVAM 189
 QY 278 QLYYFISLAFFGFLIPFVLIIYCYAA---IIRTLNAYD-HRWLWVVKASLLILVITFC 332
 DB 190 WAVEFLFTIFILLPFPVITVACYTATILKURTEAHGREQRRRAVGLAAVVLLAFVTC 249
 QY 333 FAPSNIIILIH-HANYVYNNNTDGLYFIYLIATCLASLNSCLDPFFLYFLMSK 382
 DB 250 FAPNPFVLLAHIVSRIFYCKS--YHYVVKLTCLCLNCLNCLDPFFVYFASR 298

RESULT 39
 AAB62285
 ID AAB62285 standard; Protein; 359 AA.
 XX
 AC AAB62285;
 XX
 DT 29-JUN-2001 (first entry)
 XX
 DE Human G-protein coupled receptor, PAUL.
 XX
 KW G-coupled protein receptor; PAUL; anti-HIV; antibacterial; antiviral;
 KW antifungal; protozoacide; cytostatic; antidiabetic; anorectic; human;
 KW antiasthmatic; antiparkinsonian; cardiac; hypertensive; osteopathic;
 KW antianginal; cerebroprotective; antiulcer; antimigraine; antiemetic;
 KW tranquilizer; nootropic; anticonvulsant; neuroleptic; vaccine.
 XX
 OS Homo sapiens.
 XX
 XX WO200125280-A1.
 XX 12-APR-2001.
 XX 03-OCT-2000; 2000WO-US27228.
 XX 06-OCT-1999; 99US-0413534.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Zhu Y, Li X, Vawter L;
 PI WPI; 2001-273570/28.
 XX N-PSDB; AAF57649.
 XX Novel G-coupled protein receptor, PAUL useful for treating diseases
 PT such as microbial infections, cancers, obesity, asthma, diabetes,
 PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer and
 PT allergy -
 XX
 XX Claim 1; Page 25; 34pp; English.
 XX This represents a human G-coupled protein receptor (designated PAUL).
 CC The PAUL polypeptide and polynucleotide are useful for treating and
 CC diagnosing infections such as bacterial, fungal, protozoan and viral
 CC infections, particularly infections caused by human immunodeficiency
 CC virus (HIV)-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia,
 CC bulimia, asthma, Parkinson's disease, acute heart failure, hypotension,
 CC hypertension, urinary retention, osteoporosis, angina pectoris, stroke,
 CC myocardial infarction, ulcers, allergies, benign prostatic hypertrophy,
 CC migraine, vomiting, psychotic and neurological disorders including
 CC anxiety, schizophrenia, manic depression, depression, delirium, dementia
 CC and severe mental retardation, and dyskinesia such as Huntington's
 CC disease or Gilles de la Tourette's syndrome. The PAUL sequences are
 CC useful as vaccines to induce an immunological response in a mammal. The
 CC PAUL polypeptide is also useful for identifying its modulators. The PAUL
 CC polynucleotide is valuable for chromosome localization studies and for
 CC tissue expression studies.
 XX
 XX Sequence 359 AA;
 XX SQ

Query Match 21.3%; Score 460; DB 22; Length 359;
 Query Match 21.3%; Score 460; DB 22; Length 359;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:50:54 ; Search time 1208.52 Seconds
(without alignments)
5454.265 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136

Sequence: 1 TLYTQHPVAGSQDIKMKIL.....AMARPLXPRRDIWEDIHAW 407

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09208629/runat_24062003.101625.12930/app.query.fasta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CGN 1.1 2013 @runat_24062003.101625.12930 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_p1n:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	965	45.2	585	17	A2820748	A2820748 2M0093H12
2	911	42.6	529	9	AA177828	AA177828 mt07E02.r
3	782.5	36.6	581	13	BM258063	BM258063 522739 MA
4	670	31.4	554	13	BM255484	BM255484 517232 MA
5	479.5	22.4	839	14	BQ897419	BQ897419 AGENCOURT
6	466.5	21.8	821	14	BM946931	BM946931 UI-M-EHOP
7	462	21.6	938	17	CNS04SUP	AL305674 Tetraodon
8	445	20.8	381	10	BB842550	BB842550 BB842550
9	443.5	20.8	765	13	BI183645	BI183645 UNL-P-FN-
10	416	19.5	801	13	BG924078	BG924078 60283635
11	415.5	19.5	624	13	BI394593	BI394593 pgp1n.pk0
12	406	19.0	685	9	AL660446	AL660446 AL660446
13	392.5	18.4	808	9	AJ456719	AJ456719 AJ456719
14	391.5	18.3	741	9	AJ453402	AJ453402 AJ453402
15	389.5	18.2	800	9	AJ452824	AJ452824 AJ452824
16	388	18.2	845	9	AJ456135	AJ456135 AJ456135
17	386	18.1	904	9	AL547762	AL547762 AL547762
18	386	18.1	931	9	AL532537	AL532537 AL532537
19	384	18.0	860	9	AJ446632	AJ446632 AJ446632
20	383	17.9	602	13	BM439733	BM439733 pgr1n.pk0
21	382.5	17.9	760	9	AJ456491	AJ456491 AJ456491
22	382.5	17.9	773	9	AJ451396	AJ451396 AJ451396
23	382.5	17.9	890	9	AJ456561	AJ456561 AJ456561
24	381.5	17.9	747	9	AJ446611	AJ446611 AJ446611
25	381	17.8	946	9	AL551903	AL551903 AL551903
26	377.5	17.7	2020	11	BC013202	BC013202 Homo sapi
27	377.5	17.7	2542	11	AK017378	AK017378 Mus muscu
28	377.5	17.7	3001	11	AK005013	AK005013 Mus muscu
29	374.5	17.5	730	9	AJ455912	AJ455912 AJ455912
30	374.5	17.5	740	9	AJ450921	AJ450921 AJ450921
31	373.5	17.5	749	9	AJ447002	AJ447002 AJ447002
32	371.5	17.4	766	9	AJ450835	AJ450835 AJ450835
33	366.5	17.2	746	9	AJ450259	AJ450259 AJ450259
34	363.5	17.0	942	14	BQ896389	BQ896389 AGENCOURT
35	360.5	16.9	495	9	AL713459	AL713459 DKF2P686N
36	360.5	16.9	798	9	AJ451155	AJ451155 AJ451155
37	356	16.7	542	9	AI942392	AI942392 wo77012.x
38	353.5	16.5	682	9	AJ453748	AJ453748 AJ453748
39	351.5	16.5	724	9	AJ447835	AJ447835 AJ447835
40	351.5	16.5	817	17	CNS03DLV	AL239260 Tetraodon
41	350.5	16.4	668	10	BE005953	BE005953 RCO-BN012
42	349	16.3	216	10	AW191803	AW191803 EST003 Su
43	348.5	16.3	531	12	BF426487	BF426487 df70d03.Y
44	347	16.2	909	17	CNS03PMF	AL254832 Tetraodon
45	346.5	16.2	1088	13	BM549799	BM549799 AGENCOURT

ALIGNMENTS

RESULT 1
AZ820748

LOCUS

DEFINITION

2M0093H12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0093H12 F, DNA sequence.

ACCESSION

AZ820748

VERSION

AZ820748.1

KEYWORDS

GSS.

SOURCE

house mouse

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 585)

AUTHORS

AZ820748 585 bp DNA linear GSS 20-FEB-2001
2M0093H12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0093H12 F, DNA sequence.

AZ820748.1 GI:12990656

GSS.

house mouse

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 585)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T., Reilly

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: rdunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0093 row: H column: 12
 Seq primer: CGTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 585.
 Location/Qualifiers
 1. .585

FEATURES source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="JUGC2M0093H12"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 153 a 168 c 117 g 147 t

BASE COUNT ORIGIN

Alignment Scores:
 Pred. No.: 2,98e-84 Length: 585
 Score: 965.00 Matches: 188
 Percent Similarity: 98.45% Conservative: 2
 Best Local Similarity: 97.41% Mismatches: 1
 Query Match: 45.18% Indels: 2
 DB: 17 Gaps: 0

US-09-208-629F-3 (1-407) x AZ820748 (1-585)

QY 40 AsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyPro 59
 Db 11 AATGTTTACAGCAACTCAGCAAGCAACCTTAACCTATTAAAGAGTTTAATGGGGTCCC 70
 QY 60 GluAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 79
 Db 71 CAAATAACCTTTGAAGAAATCCCACTTTCGACATAGAGGGCTGGACAGGCCACCA 130
 QY 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThr 99
 Db 131 ACTATAAAGCGGAGTGTCGCCGAGGACAGATTTTCACTTCCACGTGAATAATGTACC 190
 QY 100 IleGlyTyrLeuArgSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeu 119

Db 191 ATAGGATACCTCAGAGAAGTTCTTAACTACCAAGTATACCTGCCATCTACATCTCGTGTG 250
 QY 120 PheVal-ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrly 139
 Db 251 TTTGTGAGTGTGGTATACCAAGCAACATCGTACCCCTGTGGAAACTCTCTTAAAGGACCAA 310
 QY 139 sSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVa 159
 Db 311 ATCCATCAGTCTGGTTCATCTTTCACCAACTGGCCATCGCAGATCTCTTTCTGTGT 370
 QY 159 IthrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe-GlyGluV 179
 Db 371 CACACTGCCATTAAAGATCGCTACCATCTCAATGGCAACACTGGGTATTTGGCGAGG 430
 QY 179 alMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleL 199
 Db 431 TCACGTGGCGGATCACCAAGGTGTGTTTCTACGGCAACATGACTGCGGTATCTCTGATCC 490
 QY 199 eutrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysL 219
 Db 491 TCACCTGCATGGCATCAACCGTACCTGGCCACGGCTCACCTTTCACATACCAAGAGC 550
 QY 219 euProLysArgSerPheSerLeuLeuMetCysGly 230
 Db 551 TGCCCAAAACGAGCTTCTCCATGCTCATGTGTGGC 585

RESULT 2 AAL177828

LOCUS DEFINITION

AAL177828 529 bp mRNA linear EST 16-FEB-1997
 mt07f02.r1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:620379
 5' similar to TR:G663021 G663021 PROTEINASE ACTIVATED RECEPTOR 2.
 ; mRNA sequence.

ACCESSION VERSION

AAL177828.1 GI:1759090

KEYWORDS SOURCE

EST.
 house mouse.

ORGANISM

Mus musculus
 house mouse.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 529)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE JOURNAL

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

COMMENT

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:381203

Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 494.

FEATURES source

1. .529
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:620379"
 /clone_lib="Soares mouse 3NDMS"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5,"

3'; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

Alignment Scores:	4.94e-79	Length:	529
Pred. No.:	Score	Matches:	171
	911.00	Conservative:	3
Percent Similarity:	99.41%	Mismatches:	1
Best Local Similarity:	97.71%	Indels:	0
Query Match:	42.65%	Gaps:	0
DB:	9		

US-09-208-629F-3 (1-407) x AA177828 (1-529)

Qy	105	SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyVal	124
Db	3	AGTTCCTTAAGTACCAAGTGATCCTGCCATCTACATCTGCTGTTTGTGGTTGTGTGA	62
Qy	125	ProSerAsnIleValThrLeuTrpIysLeuSerLeuArgThrIySSerIleSerLeuVal	144
Db	63	CCAGCCCAACATCGTGACCGCTGTGGAAATCTCTCTTAAGGACCAATCCATCAGTCTGTGC	122
Qy	145	IlePheHisThrAsnLeuAlaIleAlaAspIleuPheCysValThrLeuProPheIys	164
Db	123	ATCTTTACACCAACCTGGCATCGAGATCTCTTTTGTGTGTACACTGCCATTTAAG	182

165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr 184
183 ATGCGCTACCATCTCAATGSCAACAACTGGGTATTTGGCAGGTCACTGTCGGATCACC 242

[illegible]

Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnIysLeuProLysArgSerPhe 224

Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTy²²⁵MetLeuProPheVal 244

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
85

DB 423 ATCTTGAAGCAGGAGTACCACTCTGTCCTCCGAGATCAACCACCGCCACGATGTCGT 483

QY 265 AspAlaCysGluSerProSerSerPheArgPheTyrTyrPheVal 279

Db 483 GACGCGTGGAGTCCCCCATCATCTTCCGATTCTACTACTTCGTC 527

RESULT 3

LOCUS	BM258063	581 bp	mRNA	linear	EST 17-DEC-2001
DEFINITION	522739 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.				
ACCESSION	BM258063				

VERSION	DMZ50003.1	GT.1.1/655662
KEYWORDS	EST.	
SOURCE	COW.	
ORGANISM	Bos taurus	

REFERENCES
1. Babcock, J. S. 1951. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos. 1, 1389, 1, 50, 591.

AUTHORS
Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T., Casas, E., Wray, J. E., White, J., Cho, J., Fahrrenkrug, S. C., Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chittko-McKown, C. G.,

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

TITLE	Keele, J. W. Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 526-630 (2001)
MEDLINE	21180013
COMMENT	Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18

PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG

Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1. .581

```

/db_xref="taxon:9913"
/organism="Bos taurus"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"

```

```

/lab_host="DH10B"
/note="Vector: pCMV SP
Library made from pool

```

BASE COUNT	ORIGIN	
122 a	171 c	longissimus muscle."
		117 g

Alignment Scores:	
Pred. No.:	1.99e-66
Score:	782.50
	Length
	Match

Percent Similarity:	86.08%
Best Local Similarity:	72.16%
Query Match:	36.63%
DB:	13

US-09-208-629F-3 (1-407) x BM258063 (1-581)

2 ATGGGTACCTGAGCAGCCCCCTTAAGTAC

Q1
120 TTTGCAGTAGGTATGCCGGCCAATGCGGT
Db

QY :::| | | | |:::| | | | |
140 SerIleSerLeuValIlePheHisTrpAsn
Db ACCATCCGATGACCATCTTCTACACCAAA

160 ThrLeuProPheLysIleAlaTy^rHisLeu
179 ACACTGCCCTTTAGAAATAGCTTACCATCTCT

Qy 180 MetCysArgIleThrThrValValPheTy
Db 239 ATGTGCCGGGGCCACCACGGTCATCTTCTA

Qy 200 ThrCysMetGlyIleAsnArgTyrLeuAla
299 GCCTGCATCAGTATCAACCGCTACTAGC

Qy 220 ProLysArgSerPheSerLeuLeuMetCys

Qy 240 MetLeuProPheValIleLeuLysGlnGln

```

Db      419 ATGCTCGGTTTTCTTGAAGCAGGAGTACTATCTTGTCCAGCAGACATTACCACC 478
QY      260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrThrVal 279
Db      479 TGCAGAGTCCACACACATCGAGTCTCTGTCGCCCTTCAGCTCTACTACTTCATC 538
QY      280 SerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db      539 TCCTGGCATCTCTTGATCTTAATCCCATCTTCGTCATT 580

RESULT 4
LOCUS   BM255484
DEFINITION 517232 MARC 3BOV Bos taurus cDNA 5', mRNA linear EST 17-DEC-2001
ACCESSION BM255484
VERSION   BM255484.1 GI:117891083
KEYWORDS EST.
SOURCE   cow.
ORGANISM Bos taurus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 554)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
Perteau, G., Holt, I., Karanycheva, S., Liang, F., Quackenbush, J. and
Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCGCCAGTCACGACG
Plate: 117 row: K column: 11
Seq primer: ATTAGTGACACTATAG.
FEATURES
source
Location/Qualifiers
1..554
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT 140 a 148 c 118 g 148 t
ORIGIN

Alignment Scores:
Pred. No.: 1..83e-55 Length: 554
Score: 670.00 Matches: 126
Percent Similarity: 81.42% Conservative: 23
Best Local Similarity: 68.85% Mismatches: 32
Query Match: 31.37% Indels: 2
DB: 13 Gaps: 2

US-09-208-629F-3 (1-407) x BM255484 (1-554)

QY      15 IleLysMetLysIleLeuLeuAlaAlaGlyLeuLeuPheLeuProValThrVal 34
Db      9 ATCAAGATGAGCCGCTCATCTTTGCAGCCATTGGAGCAGCTACTTCTGTCGCTGCTTCC 68

```

```

QY      35 CysGlnSerGlyIleAsnVal---SerAspAsnSerAlaLysProThrLeuThrIleLys 53
Db      69 TGTCAAGCGGCATCGAATATATCATGCACACAACTTGGCAAGCCAACTTGTCTATTAA 128
QY      54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db      129 ACCTTCCGTGGAGCTCCCAAAATTCCTTTGAAGAGTTCCCCCTTCTGCCATAGAAGC 188
QY      74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db      189 TGGACAGGAACACCAAAACTGTAATAATCCCTGGAAGAACTTGATTCAATCTC 248
QY      94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db      249 CATGTGATATATGCTACCATGGGTACTCTGACGACCCCTTTAAGTACCAATGTATACC 308
QY      114 AlaIleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTrpLys 133
Db      309 GCATCTACATCTCTGCTGTTTGCAGTAGGTATGCGGCAATGCGTGACCCCTGTGGATG 368
QY      134 LeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 153
Db      369 CTC---TTTCAAGACCAAGAACCATCGTATGACCATCTTCTACACCAACTGGCCATTGCA 425
QY      154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db      426 GACTTCTGTTTGTGTACACTGCTGCTTTAGATAGTACCATTCAATGGGAACAAC 485
QY      174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db      486 TGGGTATTGCGGAGGTCATGTGCGCGGCACACCGGTCATCTTCTATGGCAACATGTAT 545
QY      194 CysAlaIle 196
Db      546 TGCTCCATT 554

RESULT 5
LOCUS   BO897419
DEFINITION AGENCOURT_8671134 NICHD XGC Emb4 Xenopus laevis cDNA clone
ACCESSION BO897419
VERSION   BO897419.1 GI:22289433
KEYWORDS EST.
SOURCE   African Clawed frog.
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 839)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Plate: LLAM10561 row: K column: 20
High quality sequence stop: 737.
Location/Qualifiers
1..839
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4681483"
/clone_lib="NICHD XGC Emb4"
/dev_stage="embryo, stage 31-32"

```



```

Db      63 GCAGTGTCTGCTTGGAGTAAAGGTCAGAGCGCGCGTGTGTGTACATGCTGCACCTG 122
Qy      151 AlailAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsn 170
Db      123 GCATGCCGACGTCTTTCGTGTGGTCTCCCTTCAAGATCAGCTACTACTCTCTCC 182
Qy      171 GlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrValValPheTyrGly 190
Db      183 GGCACTGATGGAGTTCGGGTCTGGAATGTCCGCTTGCACCCAGCGTTTACGGG 242
Qy      191 AsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThr 210
Db      243 AACATGTAGCCTCCATCATGCTCATGAGGTATTAAGCATGACCGGTTCTTGGGGG 302
Qy      211 AlaHisPro-----PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeu 227
Db      303 GTGTATCCGATCAGTCCCTGTCTCGGCACCTCTGGCGGACCAACTTCACT----- 356
Qy      228 MetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLys 247
Db      357 ---TGGTGTGTCATTGGGTGATGGCATCATGGGGTGTGGCCCTTCTCTCAAGGAG 413
Qy      248 GlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAlaCys 267
Db      414 CAGACCACCCGAGTTCGGGACTCAACATCACCACCTGCCACGACGCTCTCAGTGAG--- 470
Qy      268 GluSerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPheGlyPheLeu 287
Db      471 AACCTGATGCAAGGCTTTTACTCGTACTACTTCTCGGCGCTTCTCCGCCATCTCTTTCT 530
Qy      288 IleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSer--- 306
Db      531 GTCCGTTGATGTTTCCACGGTCTGCTACACGTCCATCATCGGTGCTGCTGCTCTCC 590
Qy      307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeu 319
Db      591 GCGGTTGCCAACCGAGCAAGAGTCGGGGCTTG-----TTCGTGTCGGCGG--- 641
Qy      320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis 339
Db      642 ---GTGTGTGTCATCTTCATCGTCTGCTTTGGGCCCCC-AACGTCCTCTGATGTGAC 697
Qy      340 HisAlaAsnTyrTyrTyr-----HisAsnThrAspSerLeuTyrPheMetTyr 355
Db      698 -----TACCTTTTCTCTCGACAGTCTCGGTACAGAGCAGCTACTTTGCTTAC 748
Qy      356 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheVal 375
Db      749 CTCCTCTGCGTGTGTGACAGCGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
Qy      376 MetSer 377
Db      809 GCCTCC 814

RESULT 7
CNS04SUP 938 bp DNA linear GSS 24-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 001A08 of library H from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL305674.1 GI:8199451
VERSION AL305674
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 938)
AUTHORS Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bertot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,

```

```

Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 938)
Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Bertot,A., Quetier,F., Saurin,W., Bertot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
3 (bases 1 to 938)
Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1..938
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_lib="H"
/note="Genoscope sequence ID : XCOAH001BA04XD1-end : T7"
BASE COUNT 161 a 307 c 261 g 207 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 7,45e-35 Length: 938
Score: 462.00 Matches: 89
Percent Similarity: 60.08% Conservative: 57
Best Local Similarity: 36.63% Mismatches: 81
Query Match: 21.63% Indels: 17
DB: 17 Gaps: 4
US-09-208-629F-3 (1-407) x CNS04SUP (1-938)
Qy 145 llePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 23 ATCTTCATGGCGCAACCTGGCGGTGGC-GACCTGCTCTTCGTCATCTGGGTCCCGTGA 81
Qy 165 lleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 82 ATCGCTTACCCTCAAGGACGACCTGAGTCTAGGGGAGTCCCTGTGCAAGTCTCG 141
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 142 GTGGCGCTTCTTCTACGGCAACATGTACTGCTCCTCCTTCATCGGTGCATCAGCGTC 201
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 202 CAGCGCTGCAGGCGGTGCTCCACCG-----CTGCACAAGAGGGGGG 246
Qy 225 SerLeuLeuMetCysGlyIle-----ValTrpValMetValPheLeuTyrMetLeu 241
Db 247 ACGGTGGCGCGTGGGTGTCCCGGCCATCTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 306
Qy 242 PropheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 307 CCTCTGTTCTCTGATCAGCAGGTCTGTGGTGACAAACCTCGGCATCGCACCTGCCAC 366
Qy 262 AspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu 281
Db 367 GAGCTCACCAGCCCGCCAGGAGGAGGAGCGGTGCGAGC-----TACTTCTGACCATG 420
Qy 282 AlaPhePheGlyPheLeuIlePheProPheValIleIlePheCysTyrThrThrLeuIle 301
Db 421 GGAAAGCTGGGTTTATGTTCCCGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 302 HisLysLeuLysSerLys-----AspArgIleTrpLeuGlyTyrIle 315

```


AUTHORS
TITLE Caetano, A.R., Johnson, R.K. and Pomp, D.
JOURNAL Generation and sequence characterization of a normalized cDNA
COMMENT library from swine ovarian follicles
 Unpublished (2001)
 Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Fax: 402 472 6362
 Email: dpomp@unl.edu
 Oligo-dr track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message.
 Seq primer: M13 -29
 POLYA-No.

FEATURES
 source
 Location/Qualifiers
 1. .765
 /organism="Sus scrofa"
 /strain="University of Nebraska, Lincoln Swine Selection
 Lines"
 /db_xref="taxon:9823"
 /clone="UNL-P-FN-cb-d-02-0-UNL"
 /clone_lib="UNL-P-FN"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not 1; Site_2: Eco RI; The UNL-P-FN
 library is a normalized library representing porcine
 ovarian follicles, ranging between 2.0 to 10.0 mm in
 diameter, collected during 7 days of the follicular phase
 of the pig estrous cycle. This library was derived from
 the library UNL-P-F2. The tag is a string of 5-6
 nucleotides present between the Not I site and the
 oligo-dr track. The library was constructed as described
 by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
 , 1996.
 TAG SEQ=None found"
 BASE COUNT 142 a 217 c 186 g 219 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 3.69e-33 Length: 765
 Score: 443.50 Matches: 92
 Percent Similarity: 55.86% Conservative: 51
 Best Local Similarity: 35.94% Mismatches: 90
 Query Match: 20.76% Indels: 23
 DB: 13 Gaps: 7

US-09-208-629F-3 (1-407) x B1183645 (1-765)

Qy 91 SerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGln 110
 Db 9 TCAGGCTTTCCCGTGAATGAAATTTCTGCTCTACTCTAGTGGAAATGACTATGTC 68
 Qy 111 ValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThr 130
 Db 69 TTCTCCGGTGTGTACACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 128
 Qy 131 LeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsn 149
 Db 129 CTGTGGGTCTTTCTTTCCGACCGAAGAACGATCCCGCTGTGATTTACATGGCCAAT 188
 Qy 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
 Db 189 CTGGCTTGGCGGACCTCTCTGTATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 248
 Qy 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
 Db 249 CATGGCAACACTGGGTTHACGGGAATCTCTTTGCAAGTGTCTCATGCTTTTCTAC 308
 Qy 190 GlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
 Db 309 GGCAACATGTACTGCTCCATCTTCTTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 368

Qy 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
 Db 369 ATTGTGAATCCCATGGTGCAC-----CCAGGAAGAGGCAAGCTTCCCATC--- 416
 Qy 230 GlyIle-----ValTrpValMetValPheLeuTyrMetLeuProPheValIleLeu 246
 Db 417 GSGGTCCTCCCTGGGAATATGGCTGCTGATTTCTGTGGTGACCATCCCTCTGTATCGG 476
 Qy 247 LysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAla 266
 Db 477 AAGCAGACTCTCTACATCCAGCCCTTCACATCACACCTGTCTGACGCTCTTGGCCGAG 536
 Qy 267 CysGluSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPhe 286
 Db 537 ---GAGGTGTGGTAGGGACATGTTCAATATTCTCTCTCTGGCCATCGAGTCTTC 593
 Qy 287 LeuIleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSer 306
 Db 594 CTGTTCCCGCCTTCTCAGCGCGCTGCTAGCTGCTGATGATGATCAGGACGCTGCGGTCT 653
 Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLys 316
 Db 654 TCTGCCATNGACGAAACCTCAGGAAGAGGACGCGGCC-----ATTAAAG 701
 Qy 317 AlaValLeuLeuIleValIlePheThrIleCysPheAlaProThr 332
 Db 702 CTCATCAACACCGTCTGCGCATGCTGATCTGCTTACGCTTACGCTTACGCTTACGCTTAC 749

RESULT 10
 LOCUS BG924078
 DEFINITION 801 bp mRNA linear EST 05-JUN-2001
 602823635F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4952433 5',
 mRNA sequence.
 ACCESSION BG924078
 VERSION BG924078.1 GI:14304554
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10910 row: e column: 10
 High quality sequence start: 25
 High quality sequence stop: 799.
 Location/Qualifiers
 1. .801
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4952433"
 /clone_lib="NCI_CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 166 a 228 c 193 g 213 t 1 others


```

Qy 278 pheValserLeuAlaPheGlyPheLeuLeuProPheValIleIleIlePheCysTyr 297
Db 501 TTGGTCTGCCTATCGCTGGCTTTCCTTCCTGCTGCGAGTGATGCTGTTCACTTAC 560
Qy 298 ThrThrLeuIle 301
Db 561 TGCTCAGTGCTA 572

RESULT 12
AL660446
LOCUS AL660446 685 bp mRNA linear EST 13-DEC-2001
DEFINITION AL660446 XGC-neurula silurana tropicalis cdna clone TNeu043b15 5',
mRNA sequence.
ACCESSION AL660446
VERSION AL660446.1 GI:17674619
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 685)
REFERENCE Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
AUTHORS Sanger Xenopus tropicalis EST project 2001 (10_2001)
TITLE TROPICALIS_SEQUENCE ID: TNeu043b15.sp6
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinckton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE ID: TNeu043b15.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
source 1..685
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu043b15"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/notes="Vector: pCSI07; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCSI07 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 177 a 169 c 152 g 184 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 1.48e-29 Length: 685
Score: 406.00 Matches: 86
Percent Similarity: 57.75% Conservativeness: 37
Best Local Similarity: 40.38% Mismatches: 62
Query Match: 19.01% Indels: 28
DB: 7

US-09-208-629F-3 (1-407) x AL660446 (1-685)

Qy 27 LeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp----- 43
Db 74 GTCTATTCTT-----TTGTGGGGTGGCTCTGGGGCTCAAAGATGAGGATACG 124
Qy 44 -----AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyPro 59
Db 125 CTGAGAGAAATGAATCTGCTGCC-----AAGACATTTTCGGGAAAGAAA 172
Qy 60 Gln---AsnThrPheGluGluPheProLeuSerAspIleGluGlyThrPheGlyAlaThr 78
Db 173 GAAGTGGGGAATATGAGAGCTTTCCCATAGCATCTGCTGATGCTGCCAGACACCA 232
Qy 79 ThrThr-----IleLysAlaGlu-----CysProGlu 87

```

```

Db 233 ACCAATAATAATCTCTTTTCTAAGAAATCTCCAGCAATAATGCTTTAAAGTCGCCACGCGAA 292
Qy 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 293 CACACCATCACT-----AAAGTCAGCAATTCACCTTGGTGCTACTAAGTGGCAAGTA 346
Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 347 AGCAGGGAACGTATACCGGGCATCTATATCATTTGTTGCTCATTTGTTGCTGTCAGCAAT 406
Qy 128 lleValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHis 147
Db 407 GCTTTGGTACTATGATGCTGTTCCATCAGGTGACATCTGTGTCACCATCTGTCTCTAC 466
Qy 148 ThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyr 167
Db 467 GCCAGCTTAGCCACCTCTGACTGCTCTATTGCTTCATGCTGCTCTTTAAGTAGCTTAT 526
Qy 168 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValVal 187
Db 527 CACTTAAACGGCAACAACACTGGATTTTGGAGAAACCATGTGCGGGCCATGACCATTTTC 586
Qy 188. PheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyr 207
Db 587 TTGTACTTTAACAATGCTACTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
Qy 208 LeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 647 GTTGCCATTGTCATCCATTTATATACAGGAGCTTGCCG 685

RESULT 13
AJ456719
LOCUS AJ456719 808 bp mRNA linear EST 22-APR-2002
DEFINITION AJ456719 riken1 Gallus gallus CDNA clone 9f2r1, mRNA sequence.
ACCESSION AJ456719
VERSION AJ456719.1 GI:20266815
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 808)
AUTHORS Buerstedde, J.M.
TITLE Gallus gallus bursal lymphocyte EST
JOURNAL Unpublished (2002)
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
source 1..808
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="9f2r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/notes="CB inbred strain"
BASE COUNT 189 a 169 c 204 g 244 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 3.8e-28 Length: 808
Score: 392.50 Matches: 92
Percent Similarity: 48.81% Conservativeness: 51
Best Local Similarity: 31.40% Mismatches: 97
Query Match: 18.38% Indels: 53
DB: 9

US-09-208-629F-3 (1-407) x AJ456719 (1-808)

```



```

Db      506 AGGATTGAACATCAAAAGCGGTGCATATTTGCTGGATTCTAGTATTGCTCAGACA 565
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      241 LeuProPheValle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      566 CTCCTCACTCTCACTCAACCCCTATGTCAAGACGAGG-----GCTGAAGG 610
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      258 ThrThrCysHisAspValValAlaCysGluSerProSerSerPheArgPheTyrTyr 277
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      611 ATTACATGATGGAGTATCCAACTTTGGAAGAACTAAATCT-----CTTCCTGGATT 664
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      665 CTCCTGGGCGCATGTTTCATAGGATATGTACTCCACTTATATCAATCTCATCTGCTAT 724
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      725 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAAACCACTCACTGAGAACTCT 784
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCys 328
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      785 GGTGTAACAAAAGGCTCTCAACACAATATTCTTATTATTGTTGTTGTTCTCTGT 844
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      329 PheAlaProThrAsnIleLeuValIleHis 339
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      845 TTCACACCTTACCATGTGCAATATTCAACAT 877
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :

```

RESULT 18

```

AL547762      931 bp mRNA linear EST 16-FEB-2001
LOCUS      AL547762 LTI NFL006_PL2 Homo sapiens cDNA clone CS0D1017YN05 5
DEFINITION      prime, mRNA sequence.
ACCESSION      AL547762
VERSION      AL547762.1 GI:12882129
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 931)
AUTHORS      Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
   source
     1..931
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="CS0D1017YN05"
       /clone_lib="LTI NFL006_PL2"
       /tissue_type="placenta"
       /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
               was primed with a NotI-oligo(dT)-primer. Five prime end
               enriched, double-stranded cDNA was digested with Not I and
               cloned into the Not I and Eco RV sites of the pCMVSPORT 6
               vector. Library was normalized. Library was constructed by
               Life Technologies. Contact : Feng Liang Life Technologies,
               a division of Invitrogen 9800 Medical Center Drive
               Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
               Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com"

```

BASE COUNT 268 a 224 c 165 g 273 t 1 others

ALIGNMENT SCORES

```

Pred. No.: 1.96e-27 Length: 931
Score: 386.00 Matches: 80
Percent Similarity: 54.18% Conservative: 56
Best Local Similarity: 31.87% Mismatches: 95
Query Match: 18.07% Indels: 20

```

```

DB:      9      Gaps:      5
US-09-208-629F-3 (1-407) x AL547762 (1-931)
Qy      102 TyrLeuArgSerSerLeuSerThrGlnValleProAlaIleTyrIleLeuLeuPheVal 121
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      121 TATGCACATCACAGCAGCCAGGATAGTAATGCTCTGCATTACAGCTCGCTCTTCATC 180
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      181 ATTGGGCTCGTGGAACTTACTACGCTTGGTCGTCATTTTCAAAACAGGAAAAAATC 240
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      241 AACTCTACCACCTCTATTCAACAATTTGGTGATTTCTGATATATCTTTTACCACCGCT 300
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      161 LeuProPheLysIleAlaTyrHisLeuAenGlyAsnAsnTrpValPheGlyGluValMet 180
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      301 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAAATCGGAGATCGCTTG 360
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      361 TGTAGATAACTCGCCTAGTGTTTTACATCAACACATATGCGAGGTGTGAACCTTTATGACC 420
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      421 TGCCTGAGTATTGACCGCTTCATTCCTGCTGGTGCACCTCTACGCTACACAGATAAAA 480
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      481 AGGATTGAACATGCAAAAGCGGTGCATATTTGTCTGCATTCTAGTATTGCTCAGACA 540
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      541 CTCCCACTCTCATCAACCTATGTCAAAGCAGGAG-----GCTGAAAGG 585
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      258 ThrThrCysHisAspValValAlaCysGluSerProSerSerPheArgPheTyrTyr 277
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      586 ATTACATGATGGAGTATCCAAACTTTGAAGAACTAAATCT-----CTTCCTCGGATT 639
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      640 CTGCTTGGGCGCATGTTTCATAGGATATGTACTTCCACTTATATCAATCTCATCTGCTAT 699
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      700 TCTCAGATCTGCTGCAAACTTTCAGAACTGCCAAACAAAACCACTCACTGAGAAATCT 759
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCys 328
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      760 GGTGTAAACAAAAGGCTCTCAACACAATTTCTTATTATTGTTGTTGTTCTCTCT 819
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      329 PheAlaProThrAsnIleIleLeuValIleHis 339
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      820 TTCACACCTTACCATGTTGCAATATTATCAACAT 852
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 19
AJ446632      860 bp mRNA linear EST 19-APR-2002
LOCUS      AJ446632 riken1 Gallus gallus cDNA clone 14d16r1, mRNA sequence.
DEFINITION      AJ446632
ACCESSION      AJ446632
VERSION      AJ446632.1 GI:20213853
KEYWORDS      EST.
SOURCE      chicken.
ORGANISM      Gallus gallus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
               Phasianinae; Gallus.
REFERENCE      1 (bases 1 to 860)
AUTHORS      Buerstedde, J.M.
TITLE      Gallus gallus bursal lymphocyte EST
JOURNAL      Unpublished (2002)
COMMENT      Contact: Buerstedde JM

```

QY	243	PheValIleLeuYsgInGluTyrHisLeuValHisSer-GluIleThrThrCysHisAs	262
Db	642	CTGTATCTTGTCAATCAGACGGGTATATTTTCAGACCTTAAACATCACTACCTGCCATGA	701
QY	262	pValValAsp-----AlaCysGluSerPro-----	270
Db	702	TGTGTTGGCTGAAATATTGGCTCATGACATGTTCAAGTACTTTCCTCTTCACTTGC	761
QY	271	----SerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePr	289
Db	762	TTNGAGACTTTTCTAATTC-----CAGCTGNNCATCACTGGCTGGTTGCTTAC	809
QY	289	oPhe 290	
Db	810	ATAC 813	
RESULT 20			
BM439733/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			


```

SOURCE      chicken.
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 773)
AUTHORS     Buerstedde,J.M.
TITLE       Gallus gallus bursal lymphocyte EST
JOURNAL     Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinstr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
  source
    1..773
    /organism="Gallus gallus"
    /db_xref="taxon:9031"
    /clone_lib="28j6r1"
    /cell_type="bursal lymphocyte"
    /dev_stage="2-3 weeks old"
    /note="CB inbred strain"
  BASE COUNT   182 a   165 c   190 g   235 t
  ORIGIN
    Alignment Scores:
    Pred. No.:      3,41e-27      Length:      773
    Score:          382.50        Matches:     86
    Percent Similarity: 51.88%    Conservative: 52
    Best Local Similarity: 32.33% Mismatches:    82
    Query Match:    17.91%       Indels:      46
    DB:              9           Gaps:        8
  US-09-208-629F-3 (1-407) x AJ451396 (1-773)
  QY   10 AlaGlySerGlnAspLeuLysMetLys-----IleLeuLeuLeuVal 23
  Db   19 GCCGGAGGAGGAGGAGGAGGAGGATGCCGGCGCGCGGCTGTGCGTGTGCTGCTT 78
  QY   24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
  Db   79 AGCGCCTGTGCGG-----
  QY   44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
  Db   94 ---GCTGCCCTCCACAGACACAGTAGAACACAGCAGT-----
  QY   64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAla 83
  Db   130 -----TCAAAAGGAAGAAGTTTGTGTCGCGCAGAGGTTCCAGATAGT 174
  QY   84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTrpLeu 103
  Db   175 AACGCCTCTCAGAGTCATACAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 225
  QY   104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrlleLeuLeuPheValValGly 123
  Db   226 ACAGGAAACTACTACAGTTTTTCTCCACATGCTATGTCATGTCATTATCATGGT 285
  QY   124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
  Db   286 TTGCCAAGCAATGCTATGGCCATCTGGCTCTTTTTTTTTCAGAAAGAAACATCCT 345
  QY   143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
  Db   346 GCTGTGATTATATGTTAACTTGGCATTTGGCAATTAACCTGCTGTTGCTGTTGCCA 405
  QY   163 PheLysIleAlaTyrlleHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
  Db   406 CTGAGATTTCATATCATTTAAATGGCAATTAACCTGCTGTTGGGAGAGTCTCTGCAAA 465
  QY   183 IleThrThrValValPheTyrlleGlyAsnMetTyrlleCysAlaIleLeuLeuThrCysMet 202

```

```

  Db   466 GTGCTTGTGGATTTTATGGAATATGATGCTCATTTCTTTTATGACATGCTC 525
  QY   203 GlyIleAsnArgTyrlleAlaThrAlaHisProPheThrTyrlleGlnLysLeuProLysArg 222
  Db   526 AGTGTCAAAGGTATTTGGTTGTAGTGAACCCCATAGTCACATCA-----AGAAGG 576
  QY   223 SerPheSerLeuLeuMetCysGlyIle-----Val-TrpValMetValPheLeuTy 239
  Db   577 AAGTTTGAATTTGCCCTG---GGCATCTCCCTTGTGATCTTGGATCTGATTTGTGGG 633
  QY   239 rMetLeuProPheValIleLeuLysGlnGluTyrlleHisLeuValHisSerGluIleThr 259
  Db   634 AACCATTCGCGTGTATCTTGTCAATCAGACNGTGTATATTTCAGACCTTTAATCATCTAC 693
  QY   259 rCysHisAspValVal 264
  Db   694 CTGCCATGATGTTG 709
  RESULT 23
  AJ456561
  LOCUS      890 bp mRNA linear EST 22-APR-2002
  DEFINITION riken1 Gallus gallus cDNA clone 807r1, mRNA sequence.
  ACCESSION AJ456561
  VERSION AJ456561.1 GI:20266657
  KEYWORDS EST.
  SOURCE     chicken.
  ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
  REFERENCE 1 (bases 1 to 890)
  AUTHORS     Buerstedde,J.M.
  TITLE       Gallus gallus bursal lymphocyte EST
  JOURNAL     Unpublished (2002)
  COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinstr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
  FEATURES
    source
      1..890
      /organism="Gallus gallus"
      /db_xref="taxon:9031"
      /clone_lib="807r1"
      /cell_type="bursal lymphocyte"
      /dev_stage="2-3 weeks old"
      /note="CB inbred strain"
  BASE COUNT   211 a   185 c   219 g   270 t
  ORIGIN
    Alignment Scores:
    Pred. No.:      4,07e-27      Length:      890
    Score:          382.50        Matches:     84
    Percent Similarity: 51.71%    Conservative: 52
    Best Local Similarity: 31.94% Mismatches:    87
    Query Match:    17.91%       Indels:      40
    DB:              9           Gaps:        6
  US-09-208-629F-3 (1-407) x AJ456561 (1-890)
  QY   10 AlaGlySerGlnAspIleLysMetLys-----IleLeuLeuLeuVal 23
  Db   55 GCCGGAGCGAGGAGGAGGAGGATGCCGGCGCGCGGCTGTGCGTGTGCTGCTT 114
  QY   24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
  Db   115 AGCGCCTCTGTCGCG-----
  QY   44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
  Db   130 ---GCTGCCCTCCACAGACACAGTAGAACACAGCAGT-----

```

```

QY 64 GluGluPheProLeuSerAspIleGluGlyThrThrGlyAlaThrThrThrThrIleLysAla 83
Db 166 -----TCAAAGGAAGAGTTTGTGGCCAGAGGTTCCAGATACAGT 210

QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyThrLeu 103
Db 211 AACGCCTCTCAGGAGTCATACAAA-----GTGATGACTTTCAGCAAAAGTCCTT 261

QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 262 ACAGGAAATCACTACAGTTTCTTCCACAGTCTATGTCTATGTCATTCCTTTATCATTTG 321

QY 124 ValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSer 142
Db 322 TTGCCAAGCAATGCTATGGCCATCTGGGCTCTTTTTCAGAAAGAGAGAACATCCT 381

QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 382 GCTGTGATTATATGTTAACTTGGCATTGGCAGACCTTCTCTCTGTTGTTGCCA 441

QY 163 PheIleAlaTyrHisIleAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 442 CTGAAGATTGCATCATCTTAAATGGCAATAACTGCTGTTGGGAAGGTCCTCGCAA 501

QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMet 202
Db 502 GTGCTGTGTTGGATTTTATGGAATAATGATGATCTGCTCCATCTTTTATGACATGCTC 561

QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 562 AGTGTGCAAGGATTATGGGTTGTAGTGAACCCCATAGTGCACCTCAAGA---AGGAAGTCT 618

QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 619 GAAATTTGCCCTGGGCATCTCCCTTGCTATCTGATATCTGATATGTTTGTGGAAACCATTCG 678

QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSer---GluIleThrThrCysHisAs 262
Db 679 CTGTATCTGTCAATCAGACGGTGTATATTTTCAGACCTTAACATCACTACCTGCCATGA 738

QY 262 pValVal 264
Db 739 TGTGTTG 745

RESULT 24
AJ446611 747 bp mRNA linear EST 19-APR-2002
LOCUS riken1 Gallus gallus cdna clone 14c16r1, mRNA sequence.
DEFINITION
ACCESSION AJ446611
VERSION AJ446611.1 GI:20213832
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 747)
Buerstedde, J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1. .747
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="14c16r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
FEATURES
source

```

```

/dev_stage="2-3 weeks old"
/note="CB inbred strain"
BASE COUNT 178 a 158 c 187 g 224 t
ORIGIN
Alignment Scores:
Pred. NO.: 4.09e-27 Length: 747
Score: 381.50 Matches: 82
Percent Similarity: 51.35% Conservative: 51
Best Local Similarity: 31.68% Mismatches: 87
Query Match: 17.86% Indels: 39
DB: 9 Gaps: 6
US-09-208-629F-3 (1-407) x AJ446611 (1-747)
QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 18 GCCGGAGCGGAGGAGGAGGATGCCGGCGCGCGCTGTGCTGTGCTGTGCT 77
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 78 ACGGCCCTGCTGGCG-----
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 93 ---GCTGCGCTCCACAGACAGACAGTAGAACACAGCAGT-----128
QY 64 GluGluPheProLeuSerAspIleGluGlyThrThrGlyAlaThrThrThrIleLysAla 83
Db 129 -----TCAAAGGAAGAGTTTGTGGCCAGAGGTTCCAGATACAGT 173
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyThrLeu 103
Db 174 AACGCCTCTCAGGAGTCATACAAA-----GTGATGACTTTCAGCAAAAGTCCTT 224
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 225 ACAGGAAATCACTACAGTTTCTTCCACAGTCTATGTCTATGCTCTTTATCATTTG 284
QY 124 ValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSer 142
Db 285 TTGCCAAGCAATGCTATGGCCATCTGGGCTCTTTTTCAGAAAGAGAGAACATCCT 344
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 345 GCTGTGATTATATGTTAACTTGGCATTGGCAGACCTTCTCTCTGTTGTTGCCA 404
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 405 CTGAAGATTGCATCATCTTAAATGGCAATAACTGCTGTTGGGAAGGTCCTCGCAA 464
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMet 202
Db 465 GTGCTGTGTTGGATTTTATGGAATAATGATGATCTGCTCCATCTTTTATGACATGCTC 524
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 525 AGTGTGCAAGGATATGGGTTGTAGTGAACCCCATAGTGCACCTCAAGA---AGGAAGTCT 581
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 582 GAAATTTGCCCTGGGCATCTCCCTTGTATCTGATACACTGATTTGTTGGGAACCATTCG 641
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 642 CTGTATCTGTCAATCAGACGGTGTATATTTTCAGACCTTAACATCACTACCTGCCAT 698
RESULT 25
AL551903 946 bp mRNA linear EST 16-FEB-2001
LOCUS AL551903 LTI_NFL006.PL2 Homo sapiens cdna clone CS0DIO60Y018.5
DEFINITION prime, mRNA sequence.
ACCESSION AL551903

```



```

VERSION      AL551903.1  GI:12890291
KEYWORDS     EST.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 946)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 9106 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
  source
    1..946
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="CS0DI060YJ18"
      /clone_lib="LTI_NFL006_PL2"
      /tissue_type="placenta"
      /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
             was primed with a NotI-oligo(dT)-primer. Five prime end
             enriched, double-stranded cDNA was digested with Not I and
             cloned into the Not I and Eco RV sites of the pCMVSPORT 6
             vector. Library was normalized. Library was constructed by
             Life Technologies. Contact : Feng Liang Life Technologies,
             a division of Invitrogen 9800 Medical Center Drive
             Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
             Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com"
  BASE COUNT  269 a  231 c  168 g  275 t      3 others
  ORIGIN
    Alignment Scores:
    Pred. No.:      6.16e-27      Length:      946
    Score:          381.00      Matches:      80
    Percent Similarity: 53.39%      Conservative: 54
    Best Local Similarity: 31.87%      Mismatches: 97
    Query Match:    17.84%      Indels:      20
    DB:              9          Gaps:        5

  US-09-208-629f-3 (1-407) x AL551903 (1-946)

QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheVal 121
Db 127 TATGCACATCACAGCAGCCAGATAGTATGCTCTGCATTACAGCTCGTCTTCATC 186
QY 122 valGlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSerIle 141
Db 187 ATTTGGGCTCGTGGAACTTACTAGCTTGGTCTGCTCATTTGTTCAAACAGGAAAAATC 246
QY 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 247 AACTCTACCAACCCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTACCACCGCT 306
QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMet 180
Db 307 TTGCTTACAGATAGCTACTACTATGCAATGGGCTTTGATGGAGATCGAGATGGCTTG 366
QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
Db 367 TGTAGGATACTGCGVAGTGTTTTACATCAACACATATGCAGGTGTGAACCTTTATGACC 426
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 427 TGCTGTAGTATTACCGCTTCATTTGCTGTGGTGCACCCCTCTACGCTACACACAGATAAAA 486
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMet 240
Db 487 AGGATTGAACATGCAAAAGCGGTGCATATTTGCTGGATTTCTAGTATTGCTGCAGACA 546
QY 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257

```

```

Db 547 CTCCCACTCCTCATCAACCCTATGTCAAGCAGGAG-----CTGAAAGG 591
QY 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 592 ATTACATGATGATGAGATATCCAAACTTTGAGAACTAAATCT-----CTTCCTGGAT 645
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 646 CTGCTTGGGCGATGTTTCATAGGATATGTACTTCCACTTATAATCATCTCATCTGCTAT 705
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
Db 706 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCCACTCACTGAGAAATCT 765
QY 309 ArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
Db 766 GGTGTAAACAAAAGGCTCTCAACACAATATTCTTATTTATTTGTTGTTGTTCTCTCT 825
QY 329 PheAlaProThrAsnIleIleLeuValIleHis 339
Db 826 TTCACACCTTACCATGTTGCAWTTATTCAACAT 858

RESULT 26
BC013202      2020 bp      mRNA      linear      HTC 29-AUG-2001
LOCUS        Homo sapiens, clone IMAGE:4214482, mRNA.
DEFINITION   BC013202
ACCESSION    BC013202.1  GI:15342000
VERSION      HTC.
KEYWORDS     Homo sapiens.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 2020)
AUTHORS      Strausberg, R.
TITLE        Direct Submission
JOURNAL      Submitted (27-AUG-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
             Email: cgapbs@mail.nih.gov
             Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Baylor College of Medicine Human Genome
             Sequencing Center
             Center code: BCM-HGSC
             Web site: http://www.hgsc.bcm.tmc.edu/cdna/
             Contact: villalob@bcm.tmc.edu
             Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
             A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
             Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Place: 26 Row: h Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 2695873
This clone has the following problem: frame shifted.
Location/Qualifiers
  1..2020
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4214482"
    /tissue_type="Brain, anaplastic oligodendroglioma with
    1p/19q loss"
    /clone_lib="NCI CGAP_Brn67"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
  BASE COUNT  410 a  689 c  508 g  413 t

```

ORIGIN

Alignment Scores: 3.53e-26 Length: 2020
 Pred. No.: 377.50 Matches: 99
 Score: 45.43% Conservative: 60
 Percent Similarity: 28.29% Mismatches: 136
 Best Local Similarity: 17.67% Indels: 56
 Query Match: 11 Gaps: 9
 DB:

US-09-208-629f-3 (1-407) x BC013202 (1-2020)

Qy 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGly-----Gly 58
 |||||:|||||
 Db 86 TCTGACTCCAGC-----CAAGCATGAATGCCTTGAAGTGGCT 124
 |||||:|||||

Qy 59 ProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 78
 |||||:|||||
 Db 125 CCCCCAGGTCTGATCACCACTTCTCCCTGGCCAGCCAGAG-----166
 |||||:|||||

Qy 79 ThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAla 98
 |||||:|||||
 Db 167 -----CAATGGCCAGAG-----181
 |||||:|||||

Qy 99 ThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeu 118
 |||||:|||||
 Db 182 -----AGCCACTGGAGAACATGCTGTTCGCTCTTCTACCTTCTG 223
 |||||:|||||

Qy 119 LeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr 138
 |||||:|||||
 Db 224 GATTTATCTGCTTTAGTTGGCAATACCTCGCTCTGTGGCTTTTTCATCCGAGACCAC 283
 |||||:|||||

Qy 139 LysSer----IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
 |||||:|||||
 Db 284 AAGTCGGGACCCCGCCAACTGTTCTGATGATCATCTGGCGTGGCCACTTGTCTGTC 343
 |||||:|||||

Qy 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpValPheGly 177
 |||||:|||||
 Db 344 GTGCTGCTCTGCCACCCGCTGCTACCACTTCTCTGGGAACCACTGGCCATTGGG 403
 |||||:|||||

Qy 178 GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
 |||||:|||||
 Db 404 GAAATCGCATCGCTCTCACCGCTTCTCTTACCTCAACATGATGACGCGCATCTAC 463
 |||||:|||||

Qy 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln 217
 |||||:|||||
 Db 464 TTCTTACCTGATGATGACGCGGACCGTTCTCTGCGCATTTGTGACCGGTCAAGTCCCTC 523
 |||||:|||||

Qy 218 LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe 237
 |||||:|||||
 Db 524 AAGCTCGCAGGCCCTCTACGCACACCTTGTGCTTCTCTGTTGGTGGTGGCT 583
 |||||:|||||

Qy 238 LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle 257
 |||||:|||||
 Db 584 GTGCG-ATGGCCCGCTGCTGGTGAGCCACAG-----ACCGTG 621
 |||||:|||||

Qy 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
 |||||:|||||
 Db 622 CACAGCAACACACAGGTGGTCTCGCTGAGTACCGGAGAGGAGGCTCCACCATGCC 681
 |||||:|||||

Qy 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
 |||||:|||||
 Db 682 CTGGTGTCTCTGGCA---GTGGCCTTACCTTCCGTTCTATCACCGGTCACTGCTAC 738
 |||||:|||||

Qy 298 ThrThrIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
 |||||:|||||
 Db 739 CTGCTGATCATCCGACCGCTGGCGAGGCGCTGCTGTGGAGAGGCGCTCAAGACCAAG 798
 |||||:|||||

Qy 314 TyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsn 333
 |||||:|||||
 Db 799 GCAGTGGCATGATCGCATGCTGGCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 858
 |||||:|||||

Qy 334 IleIleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSer-----350
 |||||:|||||

Db 859 GTCAACCGCTCGTCTACGTCGTGCTACCGACCATCGGCGCTCTCTCGCCACCCAG 918
 |||||:|||||

Qy 351 -----LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysIleu 368
 |||||:|||||

Db 919 CGCATCTCTGGCCCTGGCAACCGCATCACCTCTCGCTCACACGCTTCAACGGGCACTC 978
 |||||:|||||

Qy 369 AspProPheLeuTyrPheValMetSerLys 378
 |||||:|||||

Db 979 GACCCCATCATGATATTTCTTCTGCTGCTGAG 1008
 |||||:|||||

RESULT 27
 AK017378
 LOCUS
 DEFINITION
 Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15.purinegic receptor P2Y, G-protein coupled 2, full insert sequence.
 AK017378
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 Mus musculus
 HTc; CAP trapper.
 AK017378.1 GI:12856588
 Mus musculus (strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone:5430432J15.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20493374
 11042159

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saiko, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saiko, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.

QY 345 Ty-HisAsnThrAspSerLeuTyPheMetTyLeuIleAlaLeuCysLeuGlySerLeu 364
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1315 TGCCACACCTCAACGCCATCAACATGGGATATACATACCCGGCGGTGCCAGCGCC 1374
 QY 365 AsnSerCysLeuAspProPheLeuTyPheValMetSerLysValVal----- 380
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1375 AACAGTTGTCTTGACCGGTACTCTACTCTTGCGAGGCGAGAGACTTGTCGGCTTTGCC 1434
 QY 381 ---AspGlnLeuAsnPro***SerAlaMetAlaArgPro----- 392
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1435 CGAGATGCCAAGCCACCGAGGAGCTTACCCCGAGCCACAGGCTCGTCGCAAGTTGGGC 1494
 QY 393 Leu**ArgProArgAspIleTrpGluAspIle 404
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1495 CTCACAGGCTAACAGAACTGTGAGGAAGATTG 1530

AK005013 3001 bp mRNA linear HTC 19-JAN-2002
 Mus musculus adult male liver cDNA, RIKEN full-length enriched
 library, clone:1300015C04:purinergic receptor P2Y, G-protein
 coupled 2, full insert sequence.
 AK005013
 AK005013.1 GI:12836638
 HTC: CAP trapper.
 Mus musculus (strain:CS7BL/6J) adult male liver cDNA to mRNA,
 clone_11b:RIKEN full-length enriched mouse cDNA library
 clone:1300015C04.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, H., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20493174
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staabli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringuet, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.
 and Hayashizaki, Y.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5 (bases 1 to 3001)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.

TITLE
 JOURNAL

Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGAGAGAGCGCCGCTACCTCGATCTTTTITTTTITTTN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGAGAGATCCAGAGCTCAATTATTAATTAACCCCCCC 3']. cDNA was
 cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
 XhoI. Host: SOLR.

FEATURES
 source

Location/Qualifiers
 1..3001
 /organism="Mus musculus"
 /strain="CS7BL/6J"
 /db_xref="FANTOM DB:1300015C04"
 /db_xref="MGD:MG1:1895057"
 /db_xref="taxon:10090"
 /clone="1300015C04"
 /sex="male"
 /tissue type="liver"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1..3001
 /gene="P2ry2"
 477..1598
 /gene="P2ry2"
 /note="data source:MGD, source key:MG1:105107,
 evidence:ISS
 purinergic receptor P2Y, G-protein coupled 2
 putative"
 /codon_start=1
 /protein_id="BAB23746.1"
 /db_xref="GI:12836639"
 /db_xref="MGD:MG1:105107"
 /translation="MAADLEPNWSTINGTWEGDELGYKCRFNEDPKYLLPVSYGVC
 VLGLCLNVALYIFLCRLKTNASTTYMFLHVASDSLVAASLPLLVYYARGDHPFS

gene
 CDS

Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

```
FEATURES
Source
1..740
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="27a5r1"
/clone_lib="rikeni"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"
BASE COUNT 176 a 158 c 184 g 219 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 1..95e-26 Length: 740
Score: 374.50 Matches: 83
Percent Similarity: 51.33% Conservative: 52
Best Local Similarity: 31.56% Mismatches: 88
Query Match: 17.53% Indels: 40
DB: 9 Gaps: 6
US-09-208-629F-3 (1-407) x AJ450921 (1-740)
QY 10 AlaGlySerGlnAspIleLysMetLys-----ileLeuileLeuVal 23
Db 19 GCCGAGCGAGGAGGAGGAGGATGCCGGCGCGCGGCTGCGCTGCTGCTGCTT 78
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 79 AGCGCCCTGCTGGCG-----
QY 44 AsnSerAlaIleProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 94 ---GTCGCGCTCCACAGACAGTAGAACACAGCAGT----- 129
QY 64 GluGluPheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAla 83
Db 130 -----TCAAAAGGAAGAAGTTCCTGGCCAGAGGTTCCAGATACACTAGT 174
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
Db 175 AACGCTCTGAGGAGGATACAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 225
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 226 ACAGGAAACTAACATACAGTTCCTCCACTGCTATGCTATGCTATGCTATGCTATGCT 285
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 286 TTGCCAAGCAATGCTATGCCCATCTGGTCTCTTTTTCAGAAACAAAGAACATCCT 345
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 346 GCTGTGATTATATGATTAACTTGGCATTGGCAGACCTTCTCTGCTGCTGCTGCTGCT 405
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArg 182
Db 406 CTGAAGATTGCATATCATTTAAATGTCATACTGCTGCTTGGGGAAGGTCTCTGCAAA 465
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMet 202
Db 466 GTGCTGTTGGATTTCCTTATGGAATAATGATGCTGCTCCATCTCTTTTATGATGCTCTC 525
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 526 AGTGTGCAAAAGGTATTTGGGTTGTAGTGAACCCCATAGTCACCTCAAGA---AGGAAGTCT 582
QY 223 SerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuPro 242
Db 583 GAAATTGCCCTGGGCATCTCCCTGCTATCTGATCTGATCTGATCTGATCTGATCTGAT 642
QY 243 PheValIleLysGlnGluTyrHisLeuValHisSerGlyIleThrThrCysHisAs 262
```

```
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 69 AGCGCCCTGCTGGCG----- 83
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 84 ---GCTGCCCTCCACAGACAGCAAGTAGAACACGAGCAGT----- 119
QY 64 GluGluPheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAla 83
Db 120 -----TCAAAAGGAAGAAGTTCCTGGCCAGAGGTTCCAGATACACTAGT 164
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
Db 165 AACGCTCTGAGGAGGATACAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 215
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 216 ACAGGAAACTAACATACAGTTCCTCCACTGCTATGCTATGCTATGCTATGCTATGCT 275
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 276 TTGCCAAGCAATGCTATGCCCATCTGGTCTCTTTTTCAGAAACAAAGAACATCCT 335
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 336 GCTGTGATTATATGATTAACTTGGCATTGGCAGACTTCTCTGCTGCTGCTGCTGCT 395
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArg 182
Db 396 CTGAAGATTGCATATCATTTAAATGGAATAACTGCTGTTGGGGAAGTCTCTGCAAA 455
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMet 202
Db 456 GTGCTTATTGG-ATTTTATGGAATAATGATGCTGCTCCATCTCTTTTATGATGATGCTC 514
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 515 AGTGTGCAAAAGGTATTTGGTGTAGTGAAACCCCATAGTCACCTCAAGA---AGGAAGTCT 571
QY 223 SerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuPro 242
Db 572 GAAATTGCCCTGGCCTCCCTGCTATCTGATGACTGATTTCCTGTTGGGAACCATCCG 631
QY 243 PheValIleLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 632 CTGTATCTTCTCAATCAGACGNGGTATATTTCAACCTTAACATCATCTACCTGCCATGAT 691
QY 263 Val 263
Db 692 GTG 694
```

RESULT 30
AJ450921 740 bp mRNA linear EST 19-APR-2002
DEFINITION AJ450921 riken1 Gallus gallus cdna clone 27a5r1, mRNA sequence.
ACCESSION AJ450921
VERSION AJ450921.1 GI:20218142
KEYWORDS EST.
SOURCE chicken.

ORGANISM
Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 740)
Buerstedde, J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany

Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de.

FEATURES

source

Location/Qualifiers
1. .495
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686N1799"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/notes="vector: pRip1Ex2; Site_1: SfiIA, Site_2: SfiIB;
cDNA-collection"
132 a 128 c 107 g 128 t

BASE COUNT 132 a 128 c 107 g 128 t

ORIGIN

Alignment Scores:
Pred. No.: 2,74e-25 Length: 495
Score: 360.50 Matches: 72
Percent Similarity: 78.90% Conservatives: 14
Best Local Similarity: 66.06% Mismatches: 22
Query Match: 16.88% Indels: 1
DB: 9 Gaps: 1

US-09-208-629F-3 (1-407) x AL713459 (1-495)

Qy 11 GlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeuPheLeu 30
Db 168 GGACTCAGTCAATCAAAATGAAGCCCTCATTTTGGAGCTGCTGGCTCTCTGTTCTG 227
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 228 TTGCCCACTTTTGTGTCAGAGTGCATGGAATATGATACAAACAACTTGGCAAGCCAAAC 287
Qy 50 LeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 288 TTACCAATTAAGACCTTTGTCGAGCTCCGCCAAATTTCTTTGAAGAGTTCCCTTTCT 347
Qy 70 AspileGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 348 GCCTTGGAGCTGGACAGAGCCAGCATTTACTGTAAATAATTAAGTCCCTGAGAAAGT 407
Qy 90 IleSerThrLeuHisValAsnAlaThrIleGlyTyLeuArgSerSerLeuSerThr 109
Db 408 GCTTCACATCTCCATGTGAAATATGTACCATGGGTACCTGACGCTCCTTAAGTACT 467
Qy 110 GlnValIleProAlaIleTyIleLeu 118
Db 468 AAAGTGTATACCTGCCATCTACCTCTG 494

RESULT 36

AJ451155 798 bp mRNA linear EST 19-APR-2002
LOCUS
DEFINITION AJ451155 riken1 Gallus gallus cDNA clone 27m8r1, mRNA sequence.
ACCESSION AJ451155
VERSION AJ451155.1 GI:20218376

KEYWORDS

SOURCE

ORGANISM

Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 798)
Buerstedde J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institut
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .798
Location/Qualifiers
/organism="Gallus gallus"

/db_xref="taxon:9031"
/clone="27m8r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"
BASE COUNT 187 a 170 c 199 g 236 t
ORIGIN

6 others

Alignment Scores:
Pred. No.: 5e-25 Length: 798
Score: 360.50 Matches: 83
Percent Similarity: 51.14% Conservatives: 52
Best Local Similarity: 31.44% Mismatches: 88
Query Match: 16.88% Indels: 41
DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x AJ451155 (1-798)

Qy 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 37 GCCGAGCGGAGGAGGAGGAGGATGCCGGCGCCGGGCTGTGCTGTGCTGTCT 96
Qy 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 97 AGCGCCCTGCTGGCG----- 111
Qy 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 112 ---GTCGCGCTCCACAGAGACAAGTAGAACACGAGT----- 147
Qy 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAla 83
Db 148 -----TCAAAGGAGAGAGTTTGTGCGCCAGAGGTTCCAGATACTAGT 192
Qy 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyLeu 103
Db 193 AACGCGCTCTGAGGAGTATACAAA-----GTGGATGACTTTGCAGCAAAAGTCCT 243
Qy 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyIleLeuLeuPheValValGly 123
Db 244 ACAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 303
Qy 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 304 TTGCCCAAGCAATGCTATGCCCATCTGGGTCTTTTTCAGAACAAAGAAACATCCT 363
Qy 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 364 GCTGTGATTATATGTTAACTTGGCATTTGGCAGACCTTCTCTGTTGTTTCCCA 423
Qy 163 PheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 424 CTGAAGATTTGCATATCATTTAAATGCAATAAATGCTGTTTGGGGAAGGTCTCTGAAA 483
Qy 183 IleThrThrValValPheTyTrpGlyAsnMetTyTrpCysAlaIleLeuLeuThrCysMet 202
Db 484 GTGCTTGTGGATTTTATGGAATATGTAATGTAATGTAATGTAATGTAATGTAATGTA 543
Qy 203 GlyIleAsnArgTyLeuAlaThrAlaHisProPheThrTyTrpGlnLysLeuProLysArg 222
Db 544 AGTGTGCAAAAGGTATTTGGTGTAGTGAACCCCATAGTGCACCTCAAGAANG---AAGTCT 600
Qy 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyTrpMetLeuPro 242
Db 601 GAAATTCCTGCGGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 660
Qy 243 PheValIleLeuLysGlnGluTyHisLeuValHis-SerGluIleThrThr-CysHisA 262
Db 661 CTGTATCTTGTCAATCANACNGNGTATATTTCAGACCTTTTAACATCACTACCTGCGCATG 720
Qy 262 spValVal 264
|||||

```

Db      721 ATGTGGTG 728

RESULT 37
AI9423392/c
LOCUS
DEFINITION
w077g12.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2461414 3'
similar to SW:PAR2_HUMAN P55085 PROTEINASE ACTIVATED RECEPTOR 2
PRECUSOR 1, mRNA sequence.
AI9423392
AI9423392.1 GI:5707048
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 542)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LUNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 932 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 356.
Location/Qualifiers
FEATURES
source
1..542
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2461414"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/notes="Organ: Kidney; Vector: pTY730-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 175 a 123 c 135 g 109 t
ORIGIN
Alignment Scores
Pred. No.: 8.44e-25 Length: 542
Score: 356.00 Matches: 74
Percent Similarity: 59.89% Conservatives: 35
Best Local Similarity: 40.66% Mismatches: 61
Query Match: 16.67% Indels: 13
DB: 9 Gaps: 4

US-09-208-629F-3 (1-407) x AI942392 (1-542)

QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValProSer 126
|||||
527 CTGACCATTCCTCTCTCCATTCACATTCGTTGCTGGTGGCTGCGCCAGT 468
|||||

QY 127 AsnIleValThr-LeuTyrPheLeuSerLeuArgThrLysSerIleSerLeu---Val11 145
|||||
467 AACGGCATGCGCTTTGGGGTCTTCTTTCCGAATTAAGAAGACACCTTCTGTGAT 408
|||||

QY 145 ePheHisThrAsnLeuAlaIleAlaPheLeuPheCysValThrLeuProPheLysI1 165
|||||

```

```

Db      407 TTACATGGCCAATCTGGCATTTGGCTGACCTCTCTGTGTCTCTGTTCCCTTGAAGAT 348

QY 165 eAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrTh 185
|||||
347 TGCCTATCACATACATACACTGGCAACAACTGGATTTATGGGAAGCTCTTTTGTAAATGCTTAT 288
|||||

QY 185 rValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAs 205
|||||
287 TGGCTTTTCTATGGCAACATCTACTGTTCCATTCTCTTCATGACCTGCTCAGTGTGCA 228
|||||

QY 205 nArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 225
|||||
227 GAGGTATTTGGTCATCGTGACCCATCGGGGACCTCCAGG---AAGAAGCAACATTGC 171
|||||

QY 225 rLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal11 245
|||||
170 CATTGGCATCTCCCTGGCAATATGCTGCTGATCTGCTGGTCACCATCCCTTTGTATGT 111
|||||

QY 245 eLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAs 265
|||||
110 CGTGAAGCAGACCATCTTTCATTCA-TGCCCTAACATCAGACCTGTTCATGATTT--- 57
|||||

QY 265 pAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSerLe 281
|||||
56 -----TCGCTTGACCACTTCTGTGGGGATCATGTTCAATTCTCTCTCT 7
|||||

QY 281 uAla 282
|||||
Db      6 GGCC 3

RESULT 38
AJ453748
LOCUS
DEFINITION
AJ453748 riken1 Gallus gallus cDNA clone 3597r1, mRNA sequence.
ACCESSION
AJ453748
VERSION
AJ453748.1 GI:20263844
KEYWORDS
EST.
SOURCE
chicken.
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 682)
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
JOURNAL
COMMENT
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
Location/Qualifiers
1..682
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="3597r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/notes="CB inbred strain"
BASE COUNT 160 a 144 c 181 g 197 t
ORIGIN
Alignment Scores
Pred. No.: 1.98e-24 Length: 682
Score: 353.50 Matches: 77
Percent Similarity: 51.22% Conservatives: 49
Best Local Similarity: 31.30% Mismatches: 81
Query Match: 16.55% Indels: 39
DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x AJ453748 (1-682)

```


RESULT 40
CNS03DLV/c
LOCUS
DEFINITION
CNS03DLV 817 bp DNA linear GSS 15-MAY-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
017G16 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION
VERSION AL239260.1 GI:7898395
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neocerigii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE
1 (bases 1 to 817)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Bounneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 817)
Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
Bounneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis.
Unpublished
3 (bases 1 to 817)
Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

Source
1..817
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="017G16"
/clone_lib="G"
/note="Genoscope sequence ID : COBG017BD08LP1-end : T7"
BASE COUNT 218 a 183 c 232 g 167 t 17 others
ORIGIN

Alignment Scores:

Pred. No.: 3.9e-24 Length: 817
Score: 351.50 Matches: 71
Percent Similarity: 59.52% Conservative: 29
Best Local Similarity: 42.26% Mismatches: 58
Query Match: 16.46% Indels: 10
DB: 17 Gaps: 3

US-09-208-629F-3 (1-407) x CNS03DLV (1-817)

QY 101 GlyTyrLeuArgSerLeuSerThrGlnValIleProAlaIle-TyrIleLeuLeuPh 120
Db 521 GGAACTCTCTGCGCGCGCCCTCACACCGTCTTCTCCCGCTGCTATACATCACCGTCTT 462
QY 120 eValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr--Ly 139
Db 461 CGTGGTGGGCTCCCAACACCCCTGGCCATCTGGGCTTCTCTTCGGCACCAGAA 402
QY 139 sSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVa 159
Db 401 GAAGCATCCATCTCCATCTTCATGGCCAACTGGCGGTGGCGGACCTGCTCTTCGTCAT 342
QY 159 lThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVa 179
Db 341 CTGGGTCCCGCTGAAAAATCGCCTACCACTCAACGGGACGACTGGGTCTACGGCGAGTC 282

QY 179 lMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeu 199
Db 281 CCTGTGCAAGGTCTCTGGTGGCCTTCTTACGGCAACATGTACTGCTCCATCGCCTTCAT 222
QY 199 uThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysIle 219
Db 221 CGCGTGCATCACGCTCCAGCGCTGCAAGCGCTGCTCCACCG--CT 177
QY 219 uProLysArgSerPheSerLeuLeuMetCysGlyIle-----ValTrpValMetVa 236
Db 176 CGACAAGAGGGGGGCGAGCGGTGCGCGCTGGCGGTGTCGGCGGCATCTGCTGTGGT 117
QY 236 lPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGl 256
Db 116 GTGGCTCATCACCGTTCCTCTGTACTGATCAGCAGGTCTCTGTGACGAGGCTCTGGTGACAAACCTGGG 57
QY 256 uIleThrThrCysHisAspVal 263
Db 56 CATCCGCACCTGCCACGAGTC 35

Search completed: June 29, 2003, 10:14:30
Job time : 1227.52 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 09:48:59 ; Search time 131.838 Seconds
(without alignments)
4582.672 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136
Sequence: 1 TLTXXHPVAGSDIKXKIL.....AMARPLXRPRIWDIHAH 407

Scoring table:
BLOSUM62 Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09208629/runat_24062003_101628_13027/app.query.fasta_1.1166
-DB=Published Applications_NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09208629 @CGN 1.1.153 @runat_24062003_101628_13027
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq*
- 11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq*
- 12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq*
- 13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq*
- 14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1443.5	67.6	1830	9 US-10-225-567A-301	Sequence 301, App
2	563	26.4	1451	9 US-10-225-567A-299	Sequence 299, App
3	509	23.8	3592	9 US-10-225-567A-323	Sequence 323, App
4	508	23.8	3299	9 US-10-208-408-24	Sequence 24, Appl

5	507.5	23.8	4895	9 US-10-225-567A-515	Sequence 515, App
6	500	23.4	1425	9 US-10-094-417-9	Sequence 9, Appl
7	500	23.4	2137	9 US-09-782-974C-75	Sequence 75, Appl
8	500	23.4	2137	9 US-10-225-567A-529	Sequence 529, App
9	496	23.2	1955	9 US-10-190-469-2	Sequence 2, Appl
10	493	23.1	1080	10 US-09-739-151-1	Sequence 1, Appl
11	471.5	22.1	2588	10 US-09-943-718-3	Sequence 3, Appl
12	458	22.1	1080	10 US-09-943-718-5	Sequence 5, Appl
13	431.5	20.2	1086	9 US-10-222-024-1	Sequence 1, Appl
14	431.5	20.2	1086	9 US-10-251-385-77	Sequence 77, Appl
15	431.5	20.2	1638	9 US-10-225-567A-111	Sequence 111, App
16	428.5	20.1	1086	9 US-10-251-385-205	Sequence 205, App
17	398.5	18.7	2070	9 US-10-225-567A-303	Sequence 303, App
18	395	18.5	1020	9 US-10-251-385-31	Sequence 31, Appl
19	389	18.5	1020	10 US-09-788-133-1	Sequence 1, Appl
20	389	18.2	1020	9 US-10-251-385-181	Sequence 181, App
21	388.5	18.2	2299	9 US-10-225-567A-224	Sequence 224, App
22	378.5	17.7	993	10 US-09-826-791-1	Sequence 1, Appl
23	378.5	17.7	1041	9 US-09-828-478-1	Sequence 1, Appl
24	378.5	17.7	1041	10 US-09-826-791-5	Sequence 5, Appl
25	378.5	17.7	1041	10 US-09-866-230-6	Sequence 6, Appl
26	378.5	17.7	1260	9 US-09-779-679-1	Sequence 1, Appl
27	378.5	17.7	1260	9 US-09-779-679-24	Sequence 24, Appl
28	378.5	17.7	1430	9 US-09-828-478-3	Sequence 3, Appl
29	378.5	17.7	1700	10 US-09-728-952-26	Sequence 26, Appl
30	378.5	17.7	2807	9 US-10-225-567A-588	Sequence 588, App
31	378	17.7	1014	9 US-10-225-567A-546	Sequence 546, App
32	376.5	17.6	2424	9 US-10-225-567A-218	Sequence 218, App
33	374	17.5	1358	9 US-10-167-192-4	Sequence 4, Appl
34	369.5	17.3	1014	9 US-10-023-775B-1	Sequence 1, Appl
35	369.5	17.3	1014	9 US-10-270-144-1	Sequence 1, Appl
36	369.5	17.3	1014	9 US-10-188-405-7	Sequence 7, Appl
37	369.5	17.3	1014	9 US-09-885-453-2	Sequence 2, Appl
38	369.5	17.3	1014	9 US-10-079-384-13	Sequence 13, Appl
39	369.5	17.3	1014	9 US-10-225-567A-646	Sequence 646, App
40	369.5	17.3	1014	10 US-09-943-798-3	Sequence 3, Appl
41	368	17.2	1301	9 US-10-024-494-7	Sequence 7, Appl
42	367	17.2	1146	9 US-10-225-567A-413	Sequence 413, App
43	366	17.1	1429	9 US-09-077-173A-1	Sequence 1, Appl
44	362.5	17.0	1098	9 US-10-225-567A-331	Sequence 331, App

ALIGNMENTS

RESULT 1
US-10-225-567A-301
; Sequence 301, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 301
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-301

Alignment Scores:
Pred. No.: 1 82e-125 Length: 1830
Score: 1443.5 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59

Query Match:	67.58%	Indels:	1
DB:	9	Gaps:	1
US-09-208-629F-3 (1-407) x US-10-225-567A-301 (1-1830)			
Qy	11	GlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeuLeuPheLeu	30
Db	127	GGGACTCAGGTGCATCAAAATGAAAGCCCTCATCTTCAGCTGCTGCCCTCTCTCTG	186
Qy	31	ProValThrValCysGlnSerGlyLe--AsnValSerAspAsnSerAlaLysProThr	49
Db	187	TTGCCCACTTTTGTTCAGATGGCATGGAAAATCATACAAACAACTTGGCAAGCCCAACC	246
Qy	50	LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer	69
Db	247	TTACCACTTAAGACCTTTCTGGGAGCTCCCCAAATTCCTTTTGAAGAGTTTCCCTCTTCT	306
Qy	70	AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer	89
Db	307	GCCTTGGAGGCTGGACAGGAGCCACGATTACTGTAAAAAATAAGTCCCTGAAGAAGT	366
Qy	90	IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrrLeuAArgSerSerLeuSerThr	109
Db	367	GCITTCACATCTCCATGTGAAAATGCTACCATGGGGTACCTGACCAGCTCCTTAAGTACT	426
Qy	110	GlnValIleProAlaIleTyrrIleLeuLeuPheValValGlyValProSerAsnIleVal	129
Db	427	AAACTGATACCTGCGCATCTACCTCCTGGTGTGTGTAGTTGGTGTCCGGGCCAATGCTGTG	486
Qy	130	ThrLeuTrpLysLeuSerLeuAArgThrLysSerIleSerLeuValIlePheHisThrAsn	149
Db	487	ACCTGTGGAGTGCTTTCTTCAGAGCCAGATCCATCTGTACCACCTGATTACTACACCAAC	546
Qy	150	LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrrHisLeu	169
Db	547	CTGGCCATTGCAGATTTCTTTTTTGTGTACATTGCCCTTAAGATAGCTTATCATCTC	606
Qy	170	AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrr	189
Db	607	AATGGGAACAACCTGGGTATTGGAGAGGTCCTGTGCGGGGCCACACAGCATCATCTCTAT	666
Qy	190	GlyAsnMetTyrrCysAlaIleLeuLeuThrCysMetGlyIleAsnAArgTyrrLeuAla	209
Db	667	GGCAACAATGACTCCTCATTCTCCTCTGCTGCATCAGCATCAACCGCTACCTGGCC	726
Qy	210	ThrAlaHisProPheThrTyrrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys	229
Db	727	ATCGTCCATCTTTTCACCTACCGGGGCTGCCCCAAGCACACCTATGCTTGGTAACATGT	786
Qy	230	GlyIleValTrpValMetValPheLeuTyrrMetLeuProPheValIleLeuLysGlnGlu	249
Db	787	GGACTGTGTGGCAACAGTTTCTTATATATGCTGTCATTTTTCATCATGAAGCAGAA	846
Qy	250	TyrrHisLeuValHisSerGluIleThrThrCysHisAspValValAlaCysGluSer	269
Db	847	TATTATCTTGTTCAGCCAGACATCACCACTGCCATGATGTTCAACACTTGGCAGTCC	906
Qy	270	ProSerSerPheArgPheTyrrTyrrPheValSerLeuAlaPhePheGlyPheLeuIlePro	289
Db	907	TCAITCTCCCTTCCAACTCTATTACTTCATCTCCTGGCATTTCTTGGATTCTTAATCCA	966
Qy	290	PheValIleIleIlePheCysTyrrThrThrLeuIleHisLysLeuLysSerLysAspArg	309
Db	967	TTTGTGCTTATCATCTACTGCTATGACGCCATCATCCGGACACTTAATGATAGATCAT	1026
Qy	310	IleTrpLeuGlyTyrrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPhe	329
Db	1027	AGATGGTGTGGATGTATTAGCCGAGTCTCTCATCTCTGTGATTTTACCATTGCTTT	1086
Qy	330	AlaProThrAsnIleLeuValIleHisHisAlaAsnTyrrTyrrHisAsnThrAsp	349
Db	1087	GCTCCAAGCAATATTATTATTATTACCATGTAACCTACTACTACTACAACAACACTGAT	1146

Qy 350 SerLeuTyrPheMetTyrLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
 Db 1147 GGCTTATATTTTATATATCTCATAGCTTTTGGCTGGGTAGTCTTAATAGTTGCTTAGAT 1206

Qy 370 PropheLeuTyrPheValMetSerLys 378
 Db 1207 CCATTCCTTTATTTTCTCAITGCAAAA 1233

RESULT 2
 US-10-225-567A-299
 ; Sequence 299, Application US/10225567A
 ; Publication No. US2003013798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LifeSpan Biosciences
 ; APPLICANT: Brown, Joseph P.
 ; APPLICANT: Burner, Glenn C.
 ; APPLICANT: Roush, Christine L.
 ; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED
 ; FILE REFERENCE: 1920-4-4
 ; CURRENT APPLICATION NUMBER: US/10/225,567A
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/257,144
 ; PRIOR FILING DATE: 2000-12-19
 ; NUMBER OF SEQ ID NOS: 2392
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 299
 ; LENGTH: 1451
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-225-567A-299

Alignment Scores:
 Pred. NO.: 3,16e-43 Length: 1451
 Score: 563.00 Matches: 119
 Percent Similarity: 52.32% Conservative: 73
 Best Local Similarity: 32.43% Mismatches: 137
 Query Match: 26.36% Indels: 38
 DB: 9 Gaps: 8

US-09-208-629F-3 (1-407) x US-10-225-567A-299 (1-1451)

Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 Db 172 CTGCTGGGGCGCCATCTCTGTAGAGCTCTCTCTCGTCAGTGCACCATCCAAGGA 231

Qy 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
 Db 232 ACCAATAGATCTCTCTAAAGGAAGAGCCTTATTGGTAAGGTTGATGCG----- 279

Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
 Db 280 -----ACATCCACGCTC-----ACTGGAAAGAGGATTACAGTT 312

Qy 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
 Db 313 GAA-----ACAGTCTTTTCTGGGATGAGTTTCTGCATCT 348

Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
 Db 349 GTCCTCACTGGAAAACATGACACGCGTCTCTCCATTTGCTACACAATTTGTTGTG 408

Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
 Db 409 GTGGGTTTGCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTCCGAACATAAGAAG 468

Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThr 160
 Db 469 CACCTGCTGTGATTATACATGGCCCAATCTGGCCTTGGCTGACCTCTCTCTGTCATCTGG 528

Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 Db 529 TTCCCTTTGAAGATTGCTATCATCATGCAACCAACTGGATTTATTTGGGAAGCTCTT 588


```
QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
Db 589 TGAATGCTGCTATTGCTTTCTATGGCAACATGACTGTCTCCATCTCTTCATGACC 648
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 649 TGCCTAGTGTGAGAGATATGGGTGATCGTGAACCCCATGGGCACTCCAGG---AAG 705
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 706 AAGCAAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
QY 241 LeuProPheValIleLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
Db 766 ATCCCTTTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825
QY 261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
Db 826 CATGATGTTTGT-----CCTGAGCAGCTCTTGGTGGGAGACATGTTTCAAT 870
QY 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
Db 871 TACTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
QY 297 TyrThrThrLeuLeuHisLysLeuLysSer-----LysAspArgIle 310
Db 931 TATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 990
QY 311 TrpLeuGlyTyrIleLysAlaValLeuLeuLeuLeuValIlePheThrIleCysPheAla 330
Db 991 AGGAAGAGGCGCCATCAACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
QY 331 ProThrAsnIleLeuLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
Db 1051 CCTAGTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
QY 351 LeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
Db 1111 GTCTATGCCCTGTACATGTAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1170
QY 371 PheLeuTyrPheValMetSer 377
Db 1171 TTGTCTATTACTTTGTTTCA 1191

RESULT 3
US-10-225-567A-323
; Sequence 323, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 323
; LENGTH: 3592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-323

Alignment Scores:
Pred. No.: 1,32e-37 Length: 3592
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
```

```
DB: 9 Gaps: 14
US-09-208-629F-3 (1-407) x US-10-225-567A-323 (1-3592)
QY 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db 360 CTGCTCTGCTGGTGGCCCGCTGCTTC-----AGTCTGTGGCGCCCGCTGTTG 404
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 405 TCTGCCCGCACCAGCCCGCCAGCCAGCAATCAAAAGCAACAAATCCACCTTAGATCCC 464
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 465 CGGTCAATTTCTTCTCAGGAACCCCAATGATGATAAATATGAACCAATTTTGGGAGGATGAGGAG 524
QY 72 GluGlyTyrThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
Db 525 AAAAATGAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGACGATCTCTTT 584
QY 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 585 CAAAAACAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 644
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 645 CTGACACTCTTTGTCCTCATCTGTGTACACCGGAGTGTGTTGTGTGCTGCTGCTGCTGCTGCT 704
QY 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 705 ATCATGGCCATCGTTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 761
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 762 TACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 822 AGCTATTACTTTTCCGCGCAGTATTGGCAGTTTGGCTCTGAATTTGTGCTGCTGCTGCTGCT 881
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn 205
Db 882 GCAGCATTTTACTGTAAACATGTACGCCCTCTATCTTGTCTCATGACGATCAAGCATGTAC 941
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
Db 942 CGGTTTCTGGCTGTGGTGTATCCCATGAGTCCCTCTCTGGGCTACTCTGGGAAGGGCT 1001
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 1002 TCCTTCACT-----TGTCTGGCCATCTGGGCTTTGGCCATCGCAGGGTAGTGCCT 1052
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 1053 CTGCTCCTCAAGGAGCAACCATCCAGGTGCCCGGCTCAACATCATCACTGCTCATGAT 1112
QY 263 ValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 1113 GTGCTCAATGAACCCCTGCTCGAAGGCTACTATGCTCTACTTCTCAGCCTTCTCTGCT 1172
QY 283 PhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeuIleHis 302
Db 1173 GTCTTC--TTTGTGCGCGTGTATCTTCCACGGTCTGTTATGTGTGCTATCATTCGA 1229
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 1230 TGTCTTAGCTCTTCCGAGTGTCCAAACCGCAGCAAGAGTCCGCGGCTTTG-----TTC 1283
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 1284 CTGTCAGCTGCT-----GTTTCTGCATCTTTCATCATTTGTTTCCGAGCCCAACAGCT 1337
QY 335 IleLeuValIleHisAlaAsnTyrTyrTyrHis-----AsnThrAspSerLeuTyr 352
```


[illegible]

```
Db 469 CTTTCAAAATCTACTACCTTCAACCGCCACCTGGTATTCGGGGTGTCTTTC 528
Qy 182 ArgileThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCys 201
Db 529 AACTGTGTACCGTGGCGCTTTACGAAACATGATTCAGCATCTCCACCATGACTGT 588
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 589 ATCAGCGTGGAGCGCTTCTGGGGTCTGTACCGCTCAGCTCCAGCGCTGGCGCGC 648
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
Db 649 CGTCTTACCGGTGGCGGTGTGCAGGACCTGCTGTCTCTCTGACCGCGCTGTC 708
Qy 242 ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHis 261
Db 709 CCGCTGGCGGCACCGATCTACCTACCGGTGCACCGCTGGGCATCATCCTGCTTC 768
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 769 GACGTCTCAAGTGCAGATGCTCCCGAGCGTGGCCATGTGGCGCTGTCTCTTACC 828
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleLeuPheCys 296
Db 829 ATCTTCATCTGCTG-----TCTCTATCCCGTTCGTGATCATCCGTGGCTGT 876
Qy 297 TyrThrThrLeuIleHisLysLeu-----LysSerLysAsp 308
Db 877 TACACGGCCACCACTCTCAAGCTGTTCGACGAGGAGCGACGCGCGGAGCAGCG 936
Qy 309 ArgileThrLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
Db 937 AGCGCGCGGTGGCGCTGGCGCGGTGTCTGCTG-----GCCTTGTACCTGTC 987
Qy 329 PheAlaProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThr 348
Db 988 TTCCGCCCCCAACAATCTCGTCTCTGGCGCACATCGTAGCGCGCTGTCTACGGC-- 1044
Qy 349 AspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 1045 AAGAGTACTACCACTGTACAGCTCAGCTGTGTCTCAGTGTCTCAACAACTGTCTG 1104
Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 1105 GACCGGTGTATTACTTGTGCTCCGG 1134
```

RESULT 7

```
US-09-782-974C-75
; Sequence 75, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 411USPHN311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
```

```
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-75

Alignment Scores: 4.34e-37 Length: 2137
Pred. No.: 500.00 Matches: 113
Score: 52.12% Conservative: 59
Percent Similarity: 34.24% Mismatches: 128
Best Local Similarity: 23.41% Indels: 30
Query Match: 9 Gaps: 8
DB:

US-09-208-629F-3 (1-407) x US-09-782-974C-75 (1-2137)

Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 150 TCTGAGCTCTACCTGCTACTTCTGCGCTGCTTCT-----GCACAGACCCGGCGGAG 203
Qy 89 SerIleSerThrLeuHisVal-----AsnAsnAlaThrIleGlyTyr 102
Db 204 GACCCCTCCAGGATGAGCTCCGAAACAGACACCGCGCCGACACCGAGCTGCAGATG 263
Qy 103 LeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 122
Db 264 CTGCGGAACCGCGCATCGCGTGGCCCTGCGCGCTGTACTCGCTGTGGCGCGGTC 323
Qy 123 GlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSer 142
Db 324 AGCATCCCGGGACACTCTTCTCTGTGGTGTCTGTCGGCGCATGGGGCCACATCC 383
Qy 143 ---LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
Db 384 CCGTGTGTCATCTTCATGATCAACCTGAGCGTCACGAGCTGATGCTGGCCAGCGGTG 443
Qy 162 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys 181
Db 444 CCTTTCCAAATCTACTACCATTTGCAACCGCCACCTGGGTATTCGGGGTGTCTGTTGC 503
Qy 182 ArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCys 201
Db 504 AAGTGTGACCGTGGCGCTTTTACGCAACATGATTTCCAGCATCTTCACCATGACTGT 563
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 564 ATCAGCGTGGAGCGCTTCTGGGGTCTGTACCGCTCAGCTCCAGCGCTGGCGCGC 623
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
Db 624 CGTCTTACCGGTGGCGCGCTGTGCAGGAGACCTGCTGTCTCTCTGACCGCGCTGTC 683
Qy 242 ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHis 261
Db 684 CCCTGGCGGCACCGATCTACCTACCGGTGCACCGCTGGGCATCATCCTGCTTC 743
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 744 GACGTCTCAAGTGCAGATGCTCCCGAGCGTGGCGCATGTCTCTTCTTCACTCACC 803
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleLeuPheCys 296
Db 804 ATCTTCATCTGCTG-----TTCTCTATCCCGTTCGTGATCATCCGTGGCTGT 851
```


Alignment Scores:

```
Pred. No.: 9,04e-37 Length: 1955
Score: 496.00 Matches: 109
Percent Similarity: 51.99% Conservative: 61
Best Local Similarity: 33.33% Mismatches: 133
Query Match: 23.22% Indels: 24
DB: 9 Gaps: 7

US-09-208-629F-3 (1-407) x US-10-190-469-2 (1-1955)

Qy 69 SerAspIleGluGlyThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 222 TCTGAGCTCACTGCTACTTCTGCGCTGCTCT-----GCACAGACCCGGCGAG 275
Qy 89 SerIleSerThrLeuHisVal-----AsnAsnAlaThrIleGlyTyr 102
Db 276 GACCCCTCCAGGATGAGGTCCCGAACAGACCGCGCGGACACGCGCTGCGAGATG 335
Qy 103 LeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheValVal 122
Db 336 CTGCGGAACCGCGCATCGCGTGGCCCTGCGTGTACTCGTGTGGTGGCGCGTC 395
Qy 123 GlyValProSerAsnIleValThrLeuTyrPheLysSerLeuArgThrLysSerIleSer 142
Db 396 AGCATCCCGGCAACCTCTCTCTGTGGTGTCTGTGCGCGCGCATGGCGGCCAGATCC 455
Qy 143 ---LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
Db 456 CGGTGCTCATCTTCATGATCAACCTGAGCGTCACGAGCATGATGTCGCGCACGGTGTG 515
Qy 162 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCys 181
Db 516 CTTTCCAAATCTACTACCTATGCAACCGCGCACCTGGTATTCGGGTGTGCTTTC 575
Qy 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCys 201
Db 576 AACGTGTGACCGTGGCGCTTTTACGCAACATGTATTCCAGCATCTCTCACCATGACCTGT 635
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 636 ATCAGCTGAGCGCTTCTCTGGGGTCTCTGATCCCGCTCAGCTCCAGCGCTGGCGCGC 695
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
Db 696 CGTCTGACGGTGGCGGTGTGTCAGGACCTGGTGTCTCTCTGACCGCGCTGTC 755
Qy 242 ProPheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIleThrCysHis 261
Db 756 CCGTGGCGCGCACCGATCTCACCTACCGGTGTCAGCGCCCTGGGCATCATCATCTGCTTC 815
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 816 GACGTCTCAAGTGAGCATGCTCCCGCGGTGGCCATGTGGCGGTGTTCTCTTCCACC 875
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
Db 876 ATCTTCATCTGCTG-----TTCTCATCCGTTCGTGATCACCGTGGCTGTGT 923
Qy 297 TyrThrThrLeuIleHisLysLeuLysSerLysAspArgIleTyr----- 311
Db 924 TACACGGCCACCATCTCAAGCTTTGCGCACGGAGGAGCGCGACGCGCGGAGCAGCGG 983
Qy 312 LeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCysPheAlaPro 331
Db 984 AGCCCGCGGTGGCGCTGGCGGTGCTGTGTCGCTTGTGCTGCTTGTGCTGCTGCTGCTG 1043
Qy 332 ThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeu 351
Db 1044 AACAACTTGTGCTCTGGCGCACATCGTGAGCGCGCTGTCTTACGGC---AAGAGGTAC 1100
Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 371
Db 1101 TACCACGTGTACAAGCTCACGCTGTGTCTAGCTGCTCAACAACTGTCTGGACCGCTTT 1160
```

```
Qy 372 LeuTyrPheValMetSerLys 378
Db 1161 GTTTATTACTTTGGTCCCGG 1181

RESULT 10
US-09-739-151-1
; Sequence 1, Application US/09739151
; Patent No. US20010029032A1
; GENERAL INFORMATION:
; APPLICANT: Yuan Zhu
; APPLICANT: Xiaotong Li
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: PAUL, A G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: GP-70567-Cl
; CURRENT APPLICATION NUMBER: US/09/739,151
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/413,534
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-739-151-1

Alignment Scores:
Pred. No.: 7,32e-37 Length: 1080
Score: 493.00 Matches: 106
Percent Similarity: 53.54% Conservative: 53
Best Local Similarity: 35.69% Mismatches: 116
Query Match: 23.08% Indels: 22
DB: 10 Gaps: 6

US-09-208-629F-3 (1-407) x US-09-739-151-1 (1-1080)

Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115
Db 28 GACAAACGCGAGCTGCGAGATGTCGGAACCCCGCGATCGCGTGGCCCTGCGCGTGGTG 87
Qy 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeuSer 135
Db 88 TACTCGCTGGTGGCGCGGTGACATCCCGGCAACCTCTCTCTCTGTTGGTGGTGTGTC 147
Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154
Db 148 CGGCGCATGGGCGGCGAGATCCCGTGGTCACTTTCATCATCACTGAGCGTCAGGAC 207
Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyr 174
Db 208 CTGATCGTGGCGGCGGTGTTGCTTTCCTCAATCTACTACCATTTGCAACCGCCACCTGG 267
Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194
Db 268 GTATTCGGGTGCTGCTTTGCAACGTTGACCGTGGCCCTTTTACGCAACATGATATCC 327
Qy 195 AlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214
Db 328 AGCATCTCACCATGACCTGATACGAGCGGTCTCTGGGGTCTCTGTACCGCTC 387
Qy 215 ThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGTCCAAAGCGTGGCGCGCGTTCGTACCGGTGGCGGTGGCGGTGGCGGTGGCGGTGG 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnIleThrHisLeuValHis 254
Db 448 CTGCTCTGACCGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 507
Qy 255 SerGluIleThrThrCysHisAspValValAlaCysGluSerProSer----- 271
```

```
Db 508 CTGGGCATCATCCTCTCGAGCTCTCAAGTGGAGCATGTCTCCACGCGTGGCCATG 567
Qy 272 -----SerPheArgPheTyTyPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
Db 568 TGGGCGGTGTCTCTTACCATCTTCATCTGCTG-----TTCTCTCATCCCG 615
Qy 290 PheVallellellePheCysTyTyThrThrLeulleHisLysLeu----- 304
Db 616 TTGGTATCATCCGTGGTGTGTACACGGCCACCATCTCAAGTGTGTGGCGCAGGAG 675
Qy 305 -----LysSerLysAspArgileTrpLeuGlyTyTyTrilleLysAlaValLeuLeu 321
Db 676 GCGCACGGCGGAGACGAGCGGCGCGTGGCGCTGGCGCGGTGTCTTGTCTG--- 732
Qy 322 LeuValillePheThrilleCysPheAlaProThrAsnillelleLeuValilleHisAla 341
Db 733 -----GCTTTGTGCACCTGCTTCCGCCCAACACTTCTGCTCTCGGCGCATCTGTG 786
Qy 342 AsnTyTyTyTyHisAsnThrAspSerLeuTyTyPheMetTyTyLeulleAlaLeuCysLeu 361
Db 787 AGCGCGCTGTCTACGGC---AGAGGCTACTACCGCTGTACAGCTCAGCTGTGTCTC 843
Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyTyPheValMetSerLys 378
Db 844 AGTGCCTCAACAACTCTGTGGACCCGTTTGTATTACTTTCGTCCTCCCG 894

RESULT 11
US-09-943-718-3
; Sequence 3, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Hufine, Constance F.
; Rossi, Devora L.
; Capone, Myriam
; Hedrick, Joseph A.
; Vicari, Alain
; Gorman, Daniel M.
; Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
; Reagents; Uses
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,718
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/009,817
; FILING DATE: 20-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0588K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1080
FEATURE:
NAME/KEY: misc_feature
LOCATION: 158
OTHER INFORMATION: /note= "residues 158, 159, and 276
probably absent, changing reading frame between those positions;
sequences provided in SEQ ID NO: 5 and 6"
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-943-718-3

Alignment Scores:
Pred. No.: 2,66e-34 Length: 2588
Score: 471.50 Matches: 119
Percent Similarity: 52.30% Conservative: 74
Best Local Similarity: 32.25% Mismatches: 155
Query Match: 22.07% Indels: 23
DB: 10 Gaps: 6

US-09-208-629f-3 (1-407) x US-09-943-718-3 (1-2588)
Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
Db 1040 CTGGAAGCGGAGGACTGCATTGCTCTCTACACTTCTGTGACTCAGC-ATCAGCGCTG 1098
Qy 42 SerAspAsnSerAlaLysProThrLeu-----ThrIleLysSerPheAsnGly----- 57
Db 1099 GCAGAGGGCATCCAGACCCCGACATCTACGATGATGTAAGAGTACCGAGGGGAAGCCAT 1158
Qy 58 -----GlyProGlnAsnThrPheGluGluPheProLeuSerAsp----- 70
Db 1159 GAAGGCCCTCTGGTCCACAGTACACTCAAGGACCGGAGCTCTCAGACAGCCTAAT 1218
Qy 71 IleGluGlyTyTyThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIle 90
Db 1219 CCACGAGGCTACCGCGGCAATTC-----TGTCACACGACAGT--- 1257
Qy 91 SerThrLeuHisValAsnAsnAlaThrIleGlyTyTyLeuArgSerSerLeuSerThrGln 110
Db 1258 GACACGCTGAGCTCCCGCCAGCTCTCAAGCACTCTCTGGGGTGGGTCCCGCAAG 1317
Qy 111 ValIleProAlaIleTyTyIleLeuLeuPheValValGlyValProSerAsnIleValThr 130
Db 1318 CTGTACTCTGCCCTCTATGGCTTGTGGCTGTGGGCTGCTGCTCCCAATGGCTGGCG 1377
Qy 131 LeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePhe-HisThrAsnLe 150
Db 1378 CTGTGGGTGTGGCCACAAGGGTGCCCGCTGCCATCCACCATCTGTCTCATGGAACCT 1437
Qy 150 uAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyHisIleuAs 170
Db 1438 GGCAGTGGCTGATCTGTCTGTGGCTGTGGCTGCCACCGACGACTGGCTTACCACCTGG 1497
Qy 170 nGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyTr 190
Db 1498 TGCCACGCGCTGCCCATTTGGTGAGGCTGCTGCCGGGTGGCCACAGCTGCCCTCTATGG 1557
Qy 190 yAsnMetTyTyCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyTyLeuAlaTh 210
Db 1558 CCACATGTATGGTTTCAGTGTGTCTGCTGGCTGCAGTCAGCTTGCACAGATACCTGGCCCT 1617
Qy 210 rAlaHisProPheTyTyGlnLysLeuProLysArgSerPheSerLeuLeuMetCysG 230
Db 1618 GGTGCATCTCTTGGCGGGCCCGTGCCTCAACGCCCTCAGCTACTGAGACTCTGTTT 1677
Qy 230 yIleValTrpValMetValPheLeuTyTyMetLeuTyTyPheValIleLeuLysGlnGlu 250
Db 1678 GGTGGCGCTGGCTCTCTGCAGCCACCCCTGGCTTGGCTTCTCACTCTCGCATCGGAGAACTT 1737
Qy 250 rHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerPr 270
```



```
QY 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 190 AACTCTACCAACCCTTATTCAACAAATTTGGTGATTTCTGATATCTTTTACCACGGCT 249
QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 250 TTGCCTACACGAATAGCCTACTACTGCAATGGCTTTGACTGGAGAATCGGAGATCGCTTG 309
QY 181 CysArgIleThrValValPheThrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
Db 310 TGATAGATAACTCGCTGCTAGTTGTTTATCATCAACATATCCAGGTGGAACCTTTATGACC 369
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 370 TGCTGAGTATTGACCGCTTCATTTGCTGTGGTGCACCTCTACGCTACAACAGATAA 429
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 430 AGGATTGAACATGCAAAAGCGTGCATATTGCTGGATTTCTGGAATTTCTGATTTGCTCAGACA 489
QY 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 490 CTCCTCACTCTCAACACCTATGCTCAAGCAGAG-----GCTGAAGG 534
QY 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 535 ATTACATGCATGGAGTATCCAACTTTGAAGAACTAAATCT-----CTTCCCTGGATT 588
QY 278 PheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleLeuLeuPheCysTyr 297
Db 589 CTGCTTGGGCGATGTTTTCATAGGATATGATCTTCCACTTATAATCATCTCTCATCTGCTAT 648
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
Db 649 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCACTCACTGAGAATCT 708
QY 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
Db 709 GGTGTAAACAAAGGCTCTCAACAAATTTATTTATTATTGTTGTTGTTGTTGTTGTTGTTGTT 768
QY 329 PheAlaProThrAsnIleLeuValIleHis-----HisAlaAsn 342
Db 769 TTCACACCTTACCATGTTGCAATTTATTCACACATATGATTAAGAGCTTCTGTTCTCTAAT 828
QY 343 Tyr-----TyrTyrHisAsnThrAsp---SerLeuTyrPheMetTyrLeuIle 357
Db 829 TTCTGGAATGTAGCCAAAGACATCTGCTCCAGATTTCTCTGCACITTT-----876
QY 358 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
Db 877 ACAGTATGCTGATGAATTCATTTGCTGTCAGACCTTTTATCTACTTCTTTGCAATGT 936
QY 378 Lys 378
Db 937 AAA 939
```

RESULT 15

```
US-10-225-567A-111
; Sequence 111, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burnet, Glena C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
```

```
; SEQ ID NO 111
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-111

Alignment Scores:
Pred. No.: 7,59e-31 Length: 1638
Score: 431.50 Matches: 95
Percent Similarity: 52.82% Conservatives: 64
Best Local Similarity: 31.56% Mismatches: 107
Query Match: 20.20% Indels: 35
DB: 9 Gaps: 9

US-09-208-629F-3 (1-407) x US-10-225-567A-111 (1-1638)
```

```
QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
Db 103 TATGCACATCACAGCAGCGCCAGGATAGTAATGCCTCTGCATTACAGCCTCGTCTTCATC 162
QY 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
Db 163 ATTGGGCTCGTGGGAACTTACTAGCCTTGGTCGTCATTTGTTCAAAACAGGAAAAAATC 222
QY 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 223 AACTCTACCAACCCTCTATTCAACAAATTTGGTGATTTCTGATATCTTTTACCACCGCT 282
QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 283 TTGCTCTACAGAAATAGCCTTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCTCTG 342
QY 181 CysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
Db 343 TGATAGTAACCTCGCTAGTGTGTACATCAACATATGCAAGCTGAGGTGAACTTTATGACC 402
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 403 TGCTGAGTATTGACCGCTTCATTTGCTGGTGCAACCTCTACGCTACAACAGATAAAA 462
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 463 AGGATTGAACATGCAAAAGCGCTGTGCATATTGTTGTTGATTTCTAGTATTGCTCAGACA 522
QY 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 523 CTCCTCACTCTCATCAACCTATGTCAAAGCAGGAG-----GCTGAAGG 567
QY 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 568 ATTACATGCATGGAGTATCCAACTTTGAAGAACTAAATCT-----CTTCCCTGGATT 621
QY 278 PheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleLeuLeuPheCysTyr 297
Db 622 CTGCTTGGGCGATGTTTTCATAGGATATGATCTTCCACTTATAATCATCTCTCTCTAT 681
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
Db 682 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCACTCACTGAGAATCT 741
QY 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
Db 742 GGTGTAAACAAAGGCTCTCAACAAATTTATTTATTATTGTTGTTGTTGTTGTTGTTGTTGTT 801
QY 329 PheAlaProThrAsnIleLeuValIleHis-----HisAlaAsn 342
Db 802 TTCACACCTTACCATGTTGCAATTTATTCACATATGATTAAGAGCTTCTGTTTCTCTAAT 861
QY 343 Tyr-----TyrTyrHisAsnThrAsp---SerLeuTyrPheMetTyrLeuIle 357
Db 862 TTCCTGGAATGTAGCCAAAGACATTCGTTCCAGATTTCTCTGCACITTT-----909
QY 358 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
```

Db 910 ACAGTAGCTGATGAACCTTCAATTGCTGATGAGCCCTTTATCTACTTCTTTGCAATG 969
Qy 378 Lys 378
Db 970 AAA 972

RESULT 16

US-10-251-385-205
; Sequence 205, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 205
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-205

Alignment Scores:

Pred. No.: 8e-31 Length: 1086
Score: 428.50 Matches: 96
Percent Similarity: 51.78% Conservative: 64
Best Local Similarity: 31.07% Mismatches: 98
Query Match: 20.06% Indels: 51
DB: 9 Gaps: 10

US-09-208-629f-3 (1-407) x US-10-251-385-205 (1-1086)

Qy 102 TyrLeuArgSerSerLeuSerThrGlnValleProAlaIleTyrIleLeuLeuPheVal 121
Db 70 TATGACATCACAGCAGCAGGATAGTAAATGCTCTGCATTACAGCCCTCGTCTTCATC 129
Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
Db 130 ATTGGGCTGTGGAAACTTACTAGCCTTGCTGCTCAATTGTTCAAAACAGGAAAAATC 189
Qy 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 190 AACTCTACCACCTCTATTCAACAAATTTGGTGATTTCTGATATCTTTTACCACGGCT 249
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 250 TTGCTACACGAATAGCTACTATGCAATGGCTTTGACTGGAAATCGGAGATCGCTTG 309
Qy 181 CysArgIleThrThrValPheValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeu 200
Db 310 TGTAGGATACTGCGCTAGTGTTTTACATCAACATATGAGTGTGAACTTTATGACC 369
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 370 TGCCTGAGTATTGACCGCTTCATTGCTGTGTGTCACCTCTACGCTACACAGATAAAA 429
Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 430 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGCTGGATTTCTAGTATTTGCTCAGACA 489
Qy 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 490 CTCCCCTCTCATCAACCCCTATGTCAAAGCAGGAG-----GCTGAAAGG 534

Qy 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 535 ATTACATGATGGAGTAGTCCAAACTTTGAAGAACTTAAATCT-----CTTCCCTGGATT 588
Qy 278 PheValSerLeuAlaPhePheGlyPheLeuIlePheProPheValIleIlePheCysTyr 297
Db 589 CTGCTTGGGGCATGTTTCATAGGATATGTAATCTTCCACTTATATCAATCTCATCTCTAT 648
Qy 298 ThrThrLeuIleHisLysLeu----- 304
Db 649 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCCACTCACTGAGAATCT 708
Qy 305 -----LysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
Db 709 GGTGTAACAAAAAGCTAAAAAC-----ACAATTAATCTT 744
Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis--- 339
Db 745 ATTATGTTGTGTTGTTCTCTGTTTACACCTTACCATGTTGCAATTAATTCACATATG 804
Qy 340 -----HisAlaAsnTyr-----TyrTyrHisAsnThrAsp--- 349
Db 805 ATTAAGAAGCTTCGTTTCTCTAATTTCTGGAATGTAGCCAAAGACATTCGTTCCAGATT 864
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 865 TCTCTGCACCTT-----ACAGTATGCTGATGAACCTCAATTCATTCATCGAC 912
Qy 370 PropheLeuTyrPheValMetSerLys 378
Db 913 CCTTTATCTACTTCTTTCATGTATAA 939

RESULT 17

US-10-225-567A-303
; Sequence 303, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 303
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-303

Alignment Scores:

Pred. No.: 1.3e-27 Length: 2070
Score: 398.50 Matches: 99
Percent Similarity: 45.43% Conservative: 60
Best Local Similarity: 28.29% Mismatches: 136
Query Match: 18.66% Indels: 55
DB: 9 Gaps: 9

US-09-208-629f-3 (1-407) x US-10-225-567A-303 (1-2070)

Qy 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGly-----Gly 58
Db 141 TCTGATCCAGC-----CAAGCATGAATGCGCTTGAATGGCT 179
Qy 59 ProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 78
Db 180 CCCCCAGGCTGTATCATCAACAACTTCTCCCTGGCCACGGCAGAG----- 221

```
QY 79 ThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAla 98
Db 222 -----CAATGTGGCCAGGAG----- 236
QY 99 ThrIleGlyTyrLeuArgSerLeuSerThrGlnValIleProAlaIleTyrIleLeu 118
Db 237 -----ACGCCACTGGAGAACATGCTGTTCGCTCTCTACCTTCTG 278
QY 119 LeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr 138
Db 279 GATTTTATCTGGCTTTGATGGCAATACCTCGCTCTGTGGCTTTTCATCCGAGACCAC 338
QY 139 LysSer---IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
Db 339 AAGTCCGGGACCCGGCCCAAGGTTCCTGATGATCTGCGCGGACGACTGTGCTGC 398
QY 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpValPheGly 177
Db 399 GTGCTGTCTGCCACCGCTGGTCTACCACTTCTCTGGGAACCACTGGCCATTGGG 458
QY 178 GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
Db 459 GAAATCCGATGCGCTCTACCGGCTTCTCTTACCTCAACATGACGCGCATCTAC 518
QY 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln 217
Db 519 TTCCTACCTGCATCAGCGCGGCGTTTCTTGCCATTGTGCACCGGTCAGTCCCTC 578
QY 218 LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe 237
Db 579 AAGCTCCGAGGCGCCCTCTACGACACCTGCGCTGTGCTCTCTGTGGTGTGGTGG 638
QY 238 LeuTyrMetLeuProPheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIle 257
Db 639 GTGGCCATGCGCGCTGTGGTGAGCCACAG-----ACGGTG 677
QY 258 ThrThrCysHisAspValAlaAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 678 CAGACCAACACACAGCGGTGTCTGCTGCAGCTGTACCGGAGAGGCTCCACCATGCC 737
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
Db 738 CTGGTGCTCTGTGCA---GTGGCCTTCACCTTCCCGTTTCATCACCACGGTCACTGTCTAC 794
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 795 CTGCTGATATCCGACCGCTCGGCGAGGCTGCTGTGGAGAGCGCTCAAGACCAAG 854
QY 314 TyrIleLysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsn 333
Db 855 GCAGTGGCATGATCGGCATAGTGTGCGCATCTTCTGTGCTCTGTCGTCGCTACCA 914
QY 334 IleIleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSer----- 350
Db 915 GTCAACCGCTCCGCTACGTGCTGACATACCGCATGGGCGCTCTCTCGCCACCCAG 974
QY 351 -----LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 975 CGCATCTGGCCCTGGCAACCGCATCACCTCTGCTCTACCGCTCAACGGGGCACTC 1034
QY 369 AspProPheLeuTyrPheValMetSerLys 378
Db 1035 GACCCCATCATGTATTCTTCGTGGCTGAG 1064

RESULT 18
US-10-251-385-31
; Sequence 31, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
```

```
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-31

Alignment Scores:
Pred. No.: 9,95e-28 Length: 1020
Score: 395.00 Matches: 88
Percent Similarity: 50.00% Conservative: 53
Best Local Similarity: 31.21% Mismatches: 123
Query Match: 18.49% Indels: 18
DB: 5 Gaps: 5

US-09-208-629f-3 (1-407) x US-10-251-385-31 (1-1020)
QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer 126
Db 85 CTGGAGAACATGCTGTTCGCTCTCTTACCTTCTGGATTATCTCTGCTTGTAGTGGC 144
QY 127 AsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer---IleSerLeuValIle 145
Db 145 AATACCTCGCTGTGGCTTTTCATCCGAGACCAACAGTCCGGGACCCCGGCAACGTG 204
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 205 TTCCTGATCATCTGCGCGTGTGCGCTGTCTGCTGCTGTCTGCTGCTGCTGCTGCTG 264
QY 166 AlaTyrHisLeuAsnGlyAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 265 GTCTACCATCTCTGGGAACCACTGGCCATTGGGGAATCGCATCGCTCTCACCGGC 324
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsn 205
Db 325 TTCCTCTTCTACCTCAACATGTAGCCAGCATCTACTTCTCACCTGCATCAGCGCGCAG 384
QY 206 ArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSer 225
Db 385 CGTTTCTGCGCATTTGTGCACCGGTCAGTCCCTCAAGTCCGAGGCGCCCTCTACGCA 444
QY 226 LeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle 245
Db 445 CACCTGGCTGTGCTCTCTGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 504
QY 246 LeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAsp 265
Db 505 AGCCACAG-----ACGTGCAGACCAACCAACACACGCGTGTCTGC 543
QY 266 AlaCysGluSerProSerSerPheArgPheTyrPheValSerLeuAlaPhePheGly 285
Db 544 CTGAGCTGTACCGGAGAGGCTCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 286 PheLeuIleProPheValIleIlePheCysTyrThrThrThrLeuIleHisLysLeuLys 305
Db 601 TTCACCTTCGCTTTCATCACCACCGGTCACTGCTACTGCTGCTGCTGCTGCTGCTGCT 660
QY 306 Ser-----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIle 321
Db 661 CAGGGCTGCTGTGGAGAGCGCTCAAGACCAAGGAGTGGGATGATCGGCACCTGCGG 720
QY 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAla 341
Db 721 CTGGCATCTCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
```



```
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 103 AACTTCAGAGAGAAATTTCCCAATGTATATCTGATATATTTTCTGGGAGTCTTG 162
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 163 GGAAATGGGTGTTCATATATGTTTCTCGAGCTTATAAGAGAGTCCACATCTGTGAAC 222
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 223 GTTTTTCATGCTAAATCTGGCCATTTTCAGATCTCTGTTCATAGACAGCGTTCCTTCAG 282
QY 165 IleAlaTyrHisLeuAsnGlyValPheGlyGluValMetCysAspGlyThr 184
DB 283 GCTGACTATATCTTAGAGCTCCCAATGGATATTTGGAGACCTGGCCTGCAGATATG 342
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
DB 343 TCTTATTCCTTGTATGTCACATGTACAGCAGTATTTATTTCTGACCGTGTGAGTGT 402
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 403 GTGCGTTTCTGGCAATGTTTCCACCTTCTGATCATCTTCTTCTGAGCGGTGAGTGT 462
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 463 GCTGATGATTCGGGTCTGTTAAAGTGGAGTCCCAAGTGGAGTGGATGCAAGAC 843
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 844 AGACTGCATAAAGCTTTGGTTATACACATGCGCTTGGCAGCAGCAATGCTTCTCAAT 903
QY 370 ProPheLeuTyrPhe 374
DB 904 CCTCTGCTCTATTAC 918
```

RESULT 24
US-09-826-791-5
; Sequence 5, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: P010914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05

```
; PRIOR APPLICATION NUMBER: 60/198,367  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatenIn Ver. 2.0  
; SEQ ID NO 5
```

```
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1041)  
US-09-826-791-5
```

```
Alignment Scores: 3 59e-26 Length: 1041  
Pred. No.: 378-50 Matches: 88  
Score: 378-50  
Percent Similarity: 49.51% Conservatives: 63  
Best Local Similarity: 28.85% Mismatches: 121  
Query Match: 17.72% Indels: 33  
DB: 10 Gaps: 8
```

US-09-208-629F-3 (1-407) x US-09-826-791-5 (1-1041)

```
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
DB 55 CCAATGGCACCTTCAGCAATAAACAACAGCAGCAACTGCACAATT-----GAA 102
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 103 AACTTCAGAGAGAAATTTTCCCAATGTATATCTGATATATTTTCTGGGAGTCTTG 162
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 163 GGAAATGGGTGTTCATATATGTTTCTGAGCCTTATAAGAGTCCACATCTGTGAAC 222
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 223 GTTTTTCATGCTAAATCTGGCCATTTTCAGATCTCTGTTCATAGCAGCGTTCCTTCAG 282
QY 165 IleAlaTyrHisLeuAsnGlyValPheGlyGluValMetCysArgIleThr 184
DB 283 GCTGACTATATCTTAGAGCTCCCAATGGATATTTGGAGACCTGGCCTGCAGATATG 342
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
DB 343 TCTTATTCCTTGTATGTCACATGTACAGCAGTATTTATTTCTGACCGTGTGAGTGT 402
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 403 GTGCGTTTCTGGCAATGTTTCCACCTTCTGATCATCTTCTTCTGAGCGGTGAGTGT 462
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 463 GCTGATGATTCGGGTCTGTTAAAGTGGAGTCCCAAGTGGAGTGGATGCAAGAC 513
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 514 ATGCTCTGAGCAGTGGCTCTGAGCAGAACGCGAGTGTACATCATGCTTAGAGCTG--- 570
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 571 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 603
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
DB 604 AACTATATTGCTTGTGTGGCTGCTGCTGCCATTTTTCACCTGCAGATCTGTAT 663
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 664 CTGCTGATCATTCGGGTCTGTTAAAGTGGAGTCCCAAGTGGAGTGGGCTCGGGTCTT 723
QY 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
DB 724 CACAGGAAGGCATGACCAACCATCATCATCTGATCATCTTCTTCTGTTCTCTG 783
QY 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
DB 784 CCCTATCACACACTGAGACCGTCCATTCAGCAGCATGGAAGTGGTTTATGCAAGAC 843
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 844 AGACTGCATAAAGCTTTGGTTATACACATGCGCTTGGCAGCAGCAATGCTTCTCAAT 903
QY 370 ProPheLeuTyrPhe 374
DB 904 CCTCTGCTCTATTAC 918
```


Db 724 CACAGGAGGCACTGACCACCATCATCATCACCTTGATCATCTTCTTCTGTTCTCCTG 783
Qy 331 ProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
Db 784 CCCTATCACACACTGAGGACCGTCCACTTCAGCACATGGAAGTGGGTTTATGCAAGAC 843
Qy 350 SerLeuSerPheMetTyrLeuIleAlaCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 844 AGACTGCATAAAGCTTTGGTTATACACACTGGCTTGGCAGCAGCAATGCTGCTCAAT 903
Qy 370 ProPheLeuTyrPhe 374
Db 904 CCTCTGCTCTATTAC 918
RESULT 25
US-09-866-230-6
; Sequence 6, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US/09/866,230
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1041)
US-09-866-230-6
Alignment Scores:
Pred. No.: 3,59e+26 Length: 1041
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 10 Gaps: 8
US-09-208-629f-3 (1-407) x US-09-866-230-6 (1-1041)
Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 55 CCAATGGCCACTTCAGCAATAACACAGCAGGAAGTGCACAATT-----GAA 102
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValPro 125
Db 103 AACTTCAGAGAGAAATTTCCCAATGTATATCTGATAATATTTTCTGGGAGTCTG 162
Qy 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 163 GGAATGGGTGTCCATATGTTTCTTCAGCCTTTATAAGAGTCCACATCTGTGAAC 222
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 223 GTTTTCATGTAATCTGGCCATTTTCAGATCTCTGTTTCATAAGCAGCGCTTCCCTCAGG 282
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 283 GCTGACTATTATCTTAGAGGCTCCAATTTGAGACTTGGCCTCAGCATATG 342
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 343 TCTATTCTCTGTATGTCAACATGTACAGCAGTATTATTCTCTGACCGTCTGAGTGT 402
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB: 10 Gaps: 8

Db 403 GTGCGTTTCTGCAATGGTTCACTCCCTTTGGCTTCTGCATGTTCACAGCATCAGAGT 462
Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 463 GCTGATCCTCTGTGGGATCATATGATCCTTATC-----ATGGCTTCTCAATA 513
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db 514 ATGCTCTGAGCAGTGGCTCTGAGCAGAACGGCAGTGTACATCATCTGTAGAGCTG-- 570
Qy 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 571 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 603
Qy 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 604 AACTATATTGCTGCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 664 CTGCTGATCATCTGGGTTCTGTTAAAGTGGAGTCCAGAAATCGGGGCTGCGGTTCT 723
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 724 CACAGGAGGCACTGACACCATCATCATCATCTGATCATCTTCTTGTGTTTCTCTG 783
Qy 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
Db 784 CCTATCACACACTGAGGACCGTCCACTTGCACACATGGAAGTGGTATTATCAAGAC 843
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 844 AGACTGCATAAAGCTTTGGTTATACACACTGGCTTGGCAGCAGCAATGCTGCTCAAT 903
Qy 370 ProPheLeuTyrPhe 374
Db 904 CCTCTGCTCTATTAC 918
RESULT 26
US-09-779-679-1
; Sequence 1, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sa
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26

```
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (105)..(1142)
; NAME/KEY: misc feature
; LOCATION: (120)..(122)
; OTHER INFORMATION: Alternative ATG start site
US-09-779-679-1
Alignment Scores:
Pred. No.: 4.73e-26 Length: 1260
Score: 378.50 Matches: 88
Percent Similarity: 49.51%
Best Local Similarity: 28.85%
Query Match: 17.72%
DB: 9
Gaps: 8

US-09-208-629F-3 (1-407) x US-09-779-679-1 (1-1260)
Qy 86 ProGluAspSerIleThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 159 CCAATGGCCACCTTCAGCAATAACACAGCAGGAACTGCACATT-----GAA 206
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 207 AACTTCAAGAGAGAAATTTTCCCAATTTGATATCTGATATATTTTCTGGGAGTCTTG 266
Qy 126 SerAsnIleValThrLeuTyrLys---LeuSerLeuArgThrIlysserIleSerLeuVal 144
Db 267 GGAATGGGTGGCCATATATGTTTCTCGACGCTTATAAGAGTCCACATCTGTGAAC 326
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 327 GTTTCATGCTAAATCTGGCAATCTCAGATCTCTGTTCATAAGCAGCTTCCTTCAGG 386
Qy 165 IleAlaTyrHisLeuAenGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr 184
Db 387 GCTGACTATTATCTTAGAGGCTCCAATTTGGATATTTGGAGACCTGGCCTGCAGGATTATG 446
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
Db 447 TCTTATTCTTGTATGTCAACATGTACAGCAGTATTTATTTCTCGACCGTGTGAGTGTT 506
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 507 GTGGTTCCTGGCAATGGTTCACCCCTTTCCGGCTTCTGCATGTCCACCATCAGGAGT 566
Qy 225 SerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuProPheVal 244
Db 567 GCCTGGATCTCTCTGGGATCATATGATCTTATC-----ATGGCTTCTCTCAATA 617
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 618 ATGCTCTCGACAGTGGCTCTGACGAGACGGAGTGTCTACATCATCTTAGAGCTG--- 674
Qy 265 AspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 675 -----AATCTCTATAAAATTCGCTAAGCTGCAGACCATG 707
Qy 282 -----AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
Db 708 AACTATATTGCTTTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
```

```
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 768 CTGCTGATCATTCGGGTTCTGTAAAGTGGAGTCCAGATCGGGCTCGGGTTCT 827
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 828 CACAGGAGGCACTGACCACTCATCATCATCTCTCTCTCTCTCTCTCTCTCTCT 887
Qy 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
Db 888 CCTATCACACACTGAGGACCTCCACTTGACGACATGGAAAGTGGTTTATGCAAGAC 947
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 948 AGACTGCATAAAGCTTTGGTTATCACACTGGCTTGGCAGCAGCCATGCTTCAAT 1007
Qy 370 PropheLeuTyrPhe 374
Db 1008 CCTCTGCTCTATTAC 1022
```

RESULT 27

```
US-09-779-679-24
; Sequence 24, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sa
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-779-679-24
```

```
Alignment Scores:
Pred. No.: 4.73e-26 Length: 1260
Score: 378.50 Matches: 88
```

Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-09-779-679-24 (1-1260)

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 159 CCAATGGCCTTCAACAGCAGGAGTCCCAATT-----GAA 206
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValPro 125
DB 207 AACTTCAAGAGAGAAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGTCTTG 266
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 267 GGAATGGGTTCATATATGTTTCTGCAGCTTATAAGAGTCCACATCTGTGAAC 326
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
DB 327 GTTTTCATGCTAAATCTGGCAATTCAGATCTCTGTTCATAAGCAGCTTCCCTTCAGG 386
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 387 GCTGACTATATCTTAGAGGCTCCAATGTGATATTTGGAGACCTGGCCTGCAGGATATG 446
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 447 TCTATTCCTGTGATGTCAACATGTACAGCAGATATTTATTTCTGACCGTGTGAGTGT 506
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 507 GTGCGTTTCTGCAATGGTTTCAACCTTTTCGGCTTCTGCATGTCACCATCAGGAGT 566
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 567 GCTGATCTCTGTGGGATCATATGATCTTATC-----ATGGCTTCTCCTCAATA 617
QY 245 IleLeuLysGlnGluTyrHisLeuValHiserGluIleThrCysHisAspValVal 264
DB 618 ATGCTCTGCAGAGTGGCTCTGAGCAGACGGAGGTGTCATCATCTGTAGAGTGT--- 674
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 675 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 707
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
DB 708 AACTATATTGCTTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 768 CTGCTGATCATCGGGTCTTGTAAAGTGGAGGTCCCAAGTGGGGCTGGCGGTTTCT 827
QY 314 TyrIleLysAla-----ValLeuIleLeuValIlePheThrIleCysPheAla 330
DB 828 CACAGGAAGGCACTGACCATCATCATCATCTGATCATCTTCTTGTGTTTCTGCTG 887
QY 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
DB 888 CCCTATCACACACTGAGGACCGCTCCACTTGACGACATGGAAAGTGGGTTTATGCAAGAC 947
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 948 AGACTGATAAAGCTTTGGTTTATACACTGCTGCTTGGCAGCAGCAATGCTGCTCAAT 1007
QY 370 ProPheLeuTyrPhe 374
DB 1008 CCTGCTCTATTAC 1022

RESULT 28
US-09-828-478-3

; Sequence 3, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1430
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-828-478-3

Alignment Scores:
Pred. No.: 5,67e-26 Length: 1430
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-09-828-478-3 (1-1430)

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 307 CCAATGGCCTTCAACAGCAGGAGTCCCAATT-----GAA 354
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValPro 135
DB 355 AACTTCAAGAGAGAAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGTCTTG 414
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 415 GGAATGGGTTCATATATGTTTCTGCGAGCTTATAAGAGTCCACATCTGTGAAC 474
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 475 GTTTTCATGCTAAATCTGGCAATTCAGATCTCTGTTTCAAGCAGCTTCCCTTCAGG 534
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 535 GCTGACTATATCTTAGAGGCTCCAATGTGATATTTGGAGACCTGCGCTGCAGGATATG 594
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 595 TCTATTCCTGTGATGTCAACATGTACAGCATATTTATTTCTGACCGTGTGAGTGT 654
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 655 GTGCGTTTCTGGCAATGGTTTCAACCTTTTCGGCTTCTGCTGCTGCTGCTGCTGCTGCT 714
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 715 GCTGATCTCTGTGGGATCATATGATCTTATC-----ATGGCTTCTCCTCAATA 765
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
DB 766 ATGCTCTGCAGAGTGGCTCTGAGCAGACGGAGGTGTCATCATCTGTAGAGTGT--- 822
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 823 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 855
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297

Db 856 AACTATATTCCTGGTGGTGGCTGCTGCTGCCATTTTTCACACTCAGACTCTGTAT 915
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 916 CTGCTGATCATTCGGTCTTGTAAAGTGGAGTCCCAAGATCGGGGCTGGGGTTTCT 975
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 976 CACAGGAAGGACCTGACCACCATCATCATCACCTTGATCATCTTCTTGTGTCTTCTG 1035
Qy 331 ProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHis----AsnThrAsp 349
Db 1036 CCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGTTTATGCAAGAC 1095
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1096 AGACTGCATAAAGCTTTGGTTATACACACTGGCTTGGCAGACCAATGCTGCTCAAT 1155
Qy 370 ProPheLeuTyrPhe 374
Db 1156 CCTCTGCTCTATTAC 1170

RESULT 29

US-09-728-952-26
; Sequence 26, Application US/09728952
; Patent No. US20020111302A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 26
; LENGTH: 1700
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)..(1272)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1700)
; OTHER INFORMATION: n = a,t,c or g
US-09-728-952-26

Alignment Scores:
Pred. No.: 7.29e-26 Length: 1700
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 10 Gaps: 8

US-09-208-629F-3 (1-407) x US-09-728-952-26 (1-1700)

Qy 86 progluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 286 CCAATGGCACCCTCAGCAATAACACAGGAGGACTGCACATT-----GAA 333
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValPro 125
Db 334 AACTTCAAGAGAGAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGTCTTG 393

Qy 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 394 GGAATGGTGTGCATATATGTTTCTGCAGCCTTATAAGAGTCCCACTGTGGAAC 453
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 454 GTTTTCATGCTAAATCTGCCATTTCCAGATCTCTCTGTTTCATAAGCACGCTTCCCTTCAGG 513
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 514 GGTGACTATTATCTTAGAGCTCCCAATTGGATATTGGAGACCTGCGCTGCAGGATATG 573
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 574 TCTTATTCCTTGTGTACATGTCACATGACACAGTATTATTCTGACCGCTGTGAGTGT 633
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLeuLeuProLysArgSerPhe 224
Db 634 GTGCGTTTCTGGCAATGGTTCACCCCTTTCGCTTCTGCATGTCTACACGATCAGGAGT 693
Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 694 GCCTGGATCCTCTGTGGGATCATATGATCCTTATC-----ATGGCTTCTCCTCAATA 744
Qy 245 IleLeuLysGlnGlnTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 745 ATGCTCTGGACAGTGGCTCTGAGCAGAGCGGAGTGTACATCATGCTTAGAGCTG--- 801
Qy 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 802 -----AATCTCTATAAAATTGCTAAGCTGCAGACCATG 834
Qy 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 835 AACTATATTGCTTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 895 CTGCTGATCATTCGGGTTCTGTTAAAGTGGAGTCCCAAGATCGGGGCTGGGGTTTCT 954
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 955 CACAGGAAGGACCTGACCACCATCATCATCACCTTGATCATCTTCTTCTGTTTCTGCTG 1014
Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHis----AsnThrAsp 349
Db 1015 CCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAAGTGGGTTTATGCAAGAC 1074
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1075 AGACTGCATAAAGCTTTGGTTATCACACTGGCTTGGCAGCAGCAATGCTGCTTCAAT 1134
Qy 370 ProPheLeuTyrPhe 374
Db 1135 CCTCTGCTCTATTAC 1149

RESULT 30

US-10-225-567A-588
; Sequence 588, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Roush, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 588
; LENGTH: 2807
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-588

Alignment Scores:

Pred. No.: 1-5e-25 Length: 2807
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: Gaps: 9

US-09-208-629F-3 (1-407) x US-10-225-567A-588 (1-2807)

```
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 318 CCAATGGCACCCTTCAGCAATAACACAGCAGAACTGCACAATT-----GAA 365
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheValValGlyValPro 125
DB 366 AACTTCAAGAGAAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGCTTG 425
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 426 GGAATGGGTGTCCATATATGTTTCTGCAGCCTTATAAGAAGTCCACATCTGTGAAC 485
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
DB 486 GTTTTCATGCTAAATCTGGCAATTTTCAGATCTCTCTTCATAGCAGCTTCCCTTCAGG 545
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 546 GCTGACTATATCTTAGAGCTCCAAATGGATATTTGGAGACCTGGCCTGCAGGATATG 605
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 606 TCTTATTCTTGTATGTACATGTACAGCAGATATTATTTCCTGACCGCTGCTGAGTGTT 665
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 666 GTGCGTGTCTGTCGAATGGTTTCAACCCCTTTTCGGCTTCTGCATGTCACCATCAGGAGT 725
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 726 GCTGATCTCTGTGGATCATATGATCTTATC-----ATGGCTTCTCAATA 776
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 284
DB 777 ATGCTCTCGACAGTGGCTCTGAGCAGACGCGCAGTGTACATCATGTCTTAGAGCTG--- 833
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 834 -----AATCTCTATAAAATTTGCTAAGCTGCAGACCATG 866
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
DB 867 AACTATATTGCTTGGTGGGCTGCTCTGCTGCCATTTTTCACATCAGCATCTGTAT 926
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 927 CTGCTGATCATTCGGGTTCGTGTTAAAGTGGAGGTCCAGAACTCGGGCTGCGGGTTCT 986
QY 314 TyrIleLysAla-----ValLeuIleLeuValIlePheThrIleCysPheAla 330
DB 987 CACAGGAAGCACTGACCAACATCATCATCACCTTGATCATCTCTTCTGTGTTTCTGCT 1046
QY 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrThrHis---AsnThrAsp 349
DB 1047 CCCTATCACACACTGAGGACCGTCCACTTGACACATGGAAAGTGGGTTTATGCAAGAC 1106
```

```
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 1107 AGACTGCATAAAGCTTTGGTTATCAGCTGGCTTGGCAGCAGCAATGCTGCTTCAAT 1166
QY 370 ProPheLeuTyrPhe 374
DB 1167 CCTCTGCTCTATTAC 1181
```

RESULT 31

US-10-225-567A-546
; Sequence 546, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 546
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-546

Alignment Scores:

Pred. No.: 3-85e-26 Length: 1014
Score: 378.00 Matches: 98
Percent Similarity: 50.83% Conservative: 55
Best Local Similarity: 32.56% Mismatches: 128
Query Match: 17.70% Indels: 20
DB: Gaps: 9

US-09-208-629F-3 (1-407) x US-10-225-567A-546 (1-1014)

```
QY 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
DB 4 GATGAACAGAGAAATCTGACAGTATCTTCTGCGACATGCCACTATGATGACTTC 63
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuPheValGlyValProSerAsn 127
DB 64 CGCAATCAAGTGATTCCACCTTGACTCTATGATCTCTGTTGTAGGCTTCTTTGGCAAT 123
QY 128 IleValThrLeuTrpLysLeuSerLeuArgThr-----LysSerIleSerLeuValIle 145
DB 124 GGCTTTGTGCTCTATGTCTCTC---ATAAAACCTATACACAAGAGTCAGCCTTCCAAGTA 180
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 155
DB 181 TACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTGCACACTGTGCTCTCCGTGTG 240
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 185
DB 241 GTCTATTATGTTTCAAAAGGCATTTGGCTCTTTGGTGACTTCTTGTGCGCGCTTCAGCACC 300
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsn 205
DB 301 TATGCTTTGTATGTCAACCTCTATTGTAGCATCTTCTTTATGACACCATGAGCTTTTTC 360
QY 206 ArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSer 225
DB 361 CGGTGATTCGAATTTGTTTTTCCAGTCCAGAACATTAATTTGTTTACACAGAAAAAGCC 420
QY 226 LeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle 245
DB 421 AGGTTTGTGTGTAGGATTTTGGATTTTGTGATTTTTCACCATTTCTCCATTTCTAATG 480
```

QY 246 LeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAsp 265
DB 481 GCCAACCACAAAAGAT-----GAGAAAAATAATACCAAGTCTTTGAGCCCCCACAA 534
QY 266 AlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGly 285
DB 535 GACAATCAAACTAAATCATATGTTTGGTCTTGCAATATGTCATTTG---TTTGTGGC 591
QY 286 PheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeu--- 304
DB 592 TTTATCATCCCTTTTGTATTATAATTTGCTGTACAAATGATATTTTGACCTTACTA 651
QY 305 ---LysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuVal 323
DB 652 AAAAATCATGAAAAAATCTGTCAAGTCATAAAAGCGCTATAGGAATGATCGTTC 711
QY 324 Ile-----PheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
DB 712 GTGACCGCTGCCTTTTGTAGTCAGTTCATGCCATATCATATTTCAAGCTACCATTCAC--- 768
QY 341 AlaAsnTyrTyrTyrHisAsn-----ThrAspSerLeuTyrPheMet----- 354
DB 769 ---CTTCATTTTACACAATGAACTAAACCTGTGATTTCTCTTAGAATGCAGAAG 825
QY 355 ---TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 373
DB 826 TCGTGGTCATAACCTTGTCTTGGCTGCATCAATTTGCTTTGACCTCTCTCTATAT 885
QY 374 Phe 374
DB 886 TTC 888

RESULT 32

US-10-225-567A-218
; Sequence 218, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glena C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-218

Alignment Scores:
Pred. No.: 1.87e-25 Length: 2424
Score: 376.50 Matches: 95
Percent Similarity: 47.44% Conservative: 72
Best Local Similarity: 26.99% Mismatches: 130
Query Match: 17.63% Indels: 55
DB: 9 Gaps: 12

US-09-208-629F-3 (1-407) x US-10-225-567A-218 (1-2424)

QY 46 AlalysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGlu 65
DB 159 GCTGTCCCAACGACGGACGGAGCGTCTCTGCGCGGTCCGGTTCGTCC----- 209
QY 66 PheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAlaGluCys 85
DB 210 -----TGGGGGAACGACGCGTCC----- 230

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAla---ThrIleGlyTyrLeuArg 104
DB 231 -----TCCACTGCCCGCTCTCTCGTTCCTCAAAATGCGCCTTGACCAAG 275
QY 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyVal 124
DB 276 ACGGCGTTCACAGTTCCTACTGCGCGTGTCTACATCTTGATTCATCATCGGCTTC 335
QY 125 ProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLys-----SerIle 141
DB 336 CTGGGCACACAGCGTGGCCATCTGGATGTCTGCTTCCCATGAAAGCCCTGGAGCGGCATC 395
QY 142 SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
DB 396 TCCGTTGATACATGTC-----AATTGGCTCTGGCGGACTTCTGTACGTGTGACTCTG 449
QY 162 PropheLysIleAlaTyrHisLeuAsnGlyAsnMetTyrPheValPheGlyGluValMetCys 181
DB 450 CCAGCGCTGATCTTCTACTTCAATAAACAAGACTGGATCTTCGGGATGCCATGTGT 509
QY 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCys 201
DB 510 AAATGCGAGAGGTTCATCTTTCATGTGAACCTCTATGGCAGCATCTTGTCTGACATGC 569
QY 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
DB 570 ATCAGTGGCCACCGGTACAGCGGTGTGTGTATCCCTCAAGTCCCTGGCGCGGTCAAA 629
QY 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
DB 630 AAGAAAGATGCGATCTGTATCAGCGTGTGTGTGCTCATTTGTGTGGTGGCGATCTCC 689
QY 242 PropheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHis 261
DB 690 CCC---ATCCTCTTCTACTCAGTACCGGGGTCCGAAAAACAAACCATCATCTGTTAC 746
QY 262 AspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeu 281
DB 747 GACACCACC-----TCAGACGAGTACCTCGGAAGTTATTTTCATCTACAGCATG 794
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThr 299
DB 795 TGCACGACCGTGGCCATGTTCTGTGCTCCCTTGGTGTGATCTGGGCTGTACCGATT 854
QY 300 LeuIleHisLysLeuLysSerLysAsp-----ArgIle 310
DB 855 ATTTGAGAGCTTTGATTTACAAAGATCTGGACAACCTCTCTCTGAGGAGAAATCGATT 914
QY 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
DB 915 TACCTG-----GTAATCATTTGACTGCTGTTTGTCTGTCTTACATC 959
QY 331 ProThrAsnIleLeuValIleHis---HisAlaAsnTyrTyrTyrHisAsn----- 347
DB 960 CTTTTCATGTGATAAAGCATGAACATTGAGGCGCGCTTGAATTTTCAGACCCCGACGA 1019
QY 348 -----ThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
DB 1020 ATGTGTGCTTCAATGACAGGCTTATGCCAGTATCAGGTGACAGAGAGGTCTAGCAAGT 1079
QY 364 LeuAsnSerCysLeuAspProPheLeuTyrPheVal 375
DB 1080 CTCAACAGTTGTGTGGACCCCATCTCTATTCTTG 1115

RESULT 33

US-10-167-192-4
; Sequence 4, Application US/10167192
; Publication No. US20030040052A1
; GENERAL INFORMATION:
; APPLICANT: Gluckmann, Maria A.
; TITLE OF INVENTION: No. US20030040052A1 G-Protein Coupled Receptors
; FILE REFERENCE: 5800-11A
; CURRENT APPLICATION NUMBER: US/10/167,192

```
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US/09/420,187
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-167-192-4

Alignment Scores:
Pred. No.: 1,39e-25 Length: 1358
Score: 374.00 Matches: 112
Percent Similarity: 47.14% Conservative: 61
Best Local Similarity: 30.52% Mismatches: 144
Query Match: 17.51% Indels: 50
DB: 14 Gaps: 14

US-09-208-629F-3 (1-407) x US-10-167-192-4 (1-1358)

QY 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 TCTGAGAACAAACGCAAGGACAGTA-----AACTGTGACCTTGAAGT 109

QY 62 ThrPheGluGluPheProLeu-----SerAspIleGluGly 73
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TAGCAGCGTGGGCTTCTCTTAATATTACACCGTAAAGGCAATTCATCACCAATAAGAAGGA 169

QY 74 TrpThrGlyAlaThrThrIleLysAla---GluCysProGlu-----87
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 -----ACATTGTGAAGGTACTCCAGTGCAGAAAGAGGACCAAAAGCAGA 214

QY 88 -----AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 215 CATTCGTAGAGAACATGGATGAACAGGAAATCTGCAGTATCTTCTGCCCATGCCAT 274

QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheVal 121
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 275 GACACTATTGATGACTTCCGCAATCAAGTGTATTCACCTTGTTACTCTATGATCTCTGTT 334

QY 122 ValCysValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr-----Lys 139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 335 GTAGGCTCTTTGGCAATGCTTTGTGCTCTATGTCCTC---ATAAAACCTATCAAG 391

QY 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 392 AAGTCAGCCTTCCAAGTATACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTC 451

QY 160 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 452 ACACTGCCTCTCCGTGTGGTCTATTATGTTCAAAAGGCATTTGGCTCTTTGGTACTTC 511

QY 180 MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeu 199
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 512 TTGTGCGCGCTCCGACCTATGCTTTGTGTGTGATGTTGATTTGTTGATTTGATTTGAT 571

QY 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 572 ACAGCCATGAGCTTTTCCCGTGCATTTGCAATTTGTTTCCAGTCCAGAACATAATTG 631

QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 632 GTTACACAGAAAAGCCAGGTTTGTGTGTGATGTTGATTTTGTGATTTGATTTGATTTG 691

QY 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThr 259
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 692 AGTTCTCCATTTCTAATGTGGCCAAACCAAAAAGAT-----GAGAAAAATAATACCAAG 745

QY 260 CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheVal 279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 746 TGCTTTGAGCCCCCACAAGACAATCAAACTAAATAATCATGTTTGGTCTTGGCAATATG 805
```



```
; APPLICANT: Tian, Hui
; APPLICANT: Dai, Kang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Zhao, Jiagang
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030082585A1el Receptors
; FILE REFERENCE: 018761-008410US
; CURRENT APPLICATION NUMBER: US/10/188, 405
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/302,800
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TGR164
US-10-188-405-7

Alignment Scores:
Pred. No.: 2,4e-25 Length: 1014
Score: 369.50 Matches: 87
Percent Similarity: 49.32% Conservative: 58
Best Local Similarity: 29.59% Mismatches: 108
Query Match: 17.30% Indels: 41
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-10-188-405-7 (1-1014)
QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer 126
Db 91 CTCAGATGACGATACCTCCCTGTTATTTATGTCATTTCTCTCGTGGATTCAGGC 150
QY 127 AsnIleValThrLeu-----TrpLysLeuSerLeuArg 137
Db 151 AATGCGATGATGATACCTTTCAAAATGAGACCTTGGAG----- 198
QY 138 ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
Db 199 -----AGCAGCACCATTCATTGCTGCACTGGCCCTGCACATCTGCTGTAT 246
QY 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 177
Db 247 CTCAGCAGCTCCCTTCCTGATTCATCTACTATGCCAGTGGCGAAACTGGATCTTGA 306
QY 178 GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
Db 307 GATTCATGTGTAAGTTTATCCGCTTCAGCTTCATTCACCTGTATAGCAGCATCCTC 366
QY 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln 217
Db 367 TTCCTCACCTGTTTCAGCATCTCCGCTACTGTGATCATTCACCAATGACCTGCTTT 426
QY 218 LysLeuProLysArgSerPheSerLeuMetCysGlyIleValTrpValMetValPhe 237
Db 427 TCATTCACAAACTCGATGTCAGTGTAGCTGTGCTGTGCTGTGATCATTTCACTG 486
QY 238 LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 487 GTAGCTGTCATTCGGATGACCTTCCTTGATCACATCAACCAAC-----AGCACCACAGA 540
QY 258 ThrThrCysHisAspValAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 541 TCAGCTGTCTCAGCTCCACCTCCAGTTCGGATGAATC---AATGACTTAAAGTGGTACAAC 597
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIlePheValIleIlePheCysTyr 297
Db 598 CTGATT---TTGACTGCAACTACTTTCGCTCCCTTGTGATGATGACACTTTGCTAT 654
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLys 307
```

```
Db 655 ACCACGATTATCCACACTCTGACCCCATGCAAACTGACAGCTGCCTTAAGCAGAAA 714
QY 308 AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIle 327
Db 715 GCACGA-----AGGCTAACCATCTCTGCTACTCTTCGATTTTACGTA 756
QY 328 CysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsn 347
Db 757 TGTITTTTACCTTCATATCTTGGAGGTCAATCGGATCGAATCTCGCCTGCTTCAATC 816
QY 348 Thr-----AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
Db 817 AGTTGTTCATTGAGAATCAGATCCATGAAGCTTACATCGTTTCTAGACCATGCTGCT 876
QY 364 LeuAsnSerCysLeuAspPhePheLeuTyrPheValMetSer 377
Db 877 CTGACACACCTTTGGTAACCTGTACTATATGTGGTGCAGC 918

RESULT 37
US-09-885-453-2
; Sequence 2, Application US/09885453
; Publication No. US2003008080A1
; GENERAL INFORMATION:
; APPLICANT: Communi, Didier
; TITLE OF INVENTION: RECEPTOR GPCRx10
; FILE REFERENCE: 9409/2082
; CURRENT APPLICATION NUMBER: US/09/885,453
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/885,453
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA nucleotide sequence
; LOCATION: (1)..(1014)
; OTHER INFORMATION: GPCRx10 DNA sequence
US-09-885-453-2

Alignment Scores:
Pred. No.: 2,4e-25 Length: 1014
Score: 369.50 Matches: 87
Percent Similarity: 49.32% Conservative: 58
Best Local Similarity: 29.59% Mismatches: 108
Query Match: 17.30% Indels: 41
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-09-885-453-2 (1-1014)
QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer 126
Db 91 CTCAGATGACGATACCTCCCTGTTATTTATGTCATTTCTCTCGTGGATTCAGGC 150
QY 127 AsnIleValThrLeu-----TrpLysLeuSerLeuArg 137
Db 151 AATGCGATGATGATACCTTTCAAAATGAGACCTTGGAG----- 198
QY 138 ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
Db 199 -----AGCAGCACCATTCATTGCTGCACTGGCCCTGCACATCTGCTGTAT 246
QY 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 177
Db 247 CTCAGCAGCTCCCTTCCTGATTCATCTACTATGCCAGTGGCGAAACTGGATCTTGA 306
QY 178 GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
Db 307 GATTCATGTGTAAGTTTATCCGCTTCAGCTTCATTCACCTGTATAGCAGCATCCTC 366
QY 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln 217
Db 367 TTCCTCACCTGTTTCAGCATCTCCGCTACTGTGATCATTCACCAATGACCTGCTTT 426
QY 218 LysLeuProLysArgSerPheSerLeuMetCysGlyIleValTrpValMetValPhe 237
Db 427 TCATTCACAAACTCGATGTCAGTGTAGCTGTGCTGTGCTGTGATCATTTCACTG 486
QY 238 LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 487 GTAGCTGTCATTCGGATGACCTTCCTTGATCACATCAACCAAC-----AGCACCACAGA 540
QY 258 ThrThrCysHisAspValAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 541 TCAGCTGTCTCAGCTCCACCTCCAGTTCGGATGAATC---AATGACTTAAAGTGGTACAAC 597
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIlePheValIleIlePheCysTyr 297
Db 598 CTGATT---TTGACTGCAACTACTTTCGCTCCCTTGTGATGATGACACTTTGCTAT 654
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLys 307
```



```

; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2392
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 646
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-646

```

Alignment Scores:		
Pred. No.:	2.4e-25	Length:
Score:	369.50	Matches:
Percent Similarity:	49.32%	Conservative:
Best Local Similarity:	29.55%	Mismatches:
Query Match:	17.30%	Indels:
DB:	9	Gaps:
		1014
		87
		58
		108
		41
		8

US-09-208-629F-3 (1-407) x US-10-225-567A-646 (1-1014)

107	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer	126	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer	127	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer	128	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer
91	CTCAAGATGCACCTAGCTCCCTGTTATTTATGGCATTTATCTCTCGTGGGATTTCCAGGC	150	CTCAAGATGCACCTAGCTCCCTGTTATTTATGGCATTTATCTCTCGTGGGATTTCCAGGC	151	CTCAAGATGCACCTAGCTCCCTGTTATTTATGGCATTTATCTCTCGTGGGATTTCCAGGC	152	CTCAAGATGCACCTAGCTCCCTGTTATTTATGGCATTTATCTCTCGTGGGATTTCCAGGC
127	AsnIleValThrLeu-----TTPlylsLeuSerLeuArg	137	AsnIleValThrLeu-----TTPlylsLeuSerLeuArg	138	AsnIleValThrLeu-----TTPlylsLeuSerLeuArg	139	AsnIleValThrLeu-----TTPlylsLeuSerLeuArg
151	AATGCAGTAGTCATATCCACTTTACATTTTCAAAATGAGACTTGAAG-----	198	AATGCAGTAGTCATATCCACTTTACATTTTCAAAATGAGACTTGAAG-----	199	AATGCAGTAGTCATATCCACTTTACATTTTCAAAATGAGACTTGAAG-----	200	AATGCAGTAGTCATATCCACTTTACATTTTCAAAATGAGACTTGAAG-----
138	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPhe	157	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPhe	158	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPhe	159	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPhe
199	-----AGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATCTGCTGTAT	246	-----AGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATCTGCTGTAT	247	-----AGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATCTGCTGTAT	248	-----AGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATCTGCTGTAT
158	CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnIlePheGly	17	CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnIlePheGly	178	CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnIlePheGly	179	CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnIlePheGly
247	CTGACGACCTCCCTTCTGATTACACTACTTGCAGATGGCGAAATCTGGATCTTTGGA	306	CTGACGACCTCCCTTCTGATTACACTACTTGCAGATGGCGAAATCTGGATCTTTGGA	307	CTGACGACCTCCCTTCTGATTACACTACTTGCAGATGGCGAAATCTGGATCTTTGGA	308	CTGACGACCTCCCTTCTGATTACACTACTTGCAGATGGCGAAATCTGGATCTTTGGA
178	GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	197	GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	198	GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	199	GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu
307	GATTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAACTGTATAGCAGCATCCCTC	366	GATTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAACTGTATAGCAGCATCCCTC	367	GATTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAACTGTATAGCAGCATCCCTC	368	GATTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAACTGTATAGCAGCATCCCTC
198	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	217	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	218	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	219	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln
367	TTTCTCACTGTTTCAGCATCTTCGCTACTGCTGTGATCATTACACCAAGAGCTGCTTT	426	TTTCTCACTGTTTCAGCATCTTCGCTACTGCTGTGATCATTACACCAAGAGCTGCTTT	427	TTTCTCACTGTTTCAGCATCTTCGCTACTGCTGTGATCATTACACCAAGAGCTGCTTT	428	TTTCTCACTGTTTCAGCATCTTCGCTACTGCTGTGATCATTACACCAAGAGCTGCTTT
218	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValThrPvalMetValPhe	237	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValThrPvalMetValPhe	238	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValThrPvalMetValPhe	239	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValThrPvalMetValPhe
427	TCCATTACAAAACCTCGATGTGCAGTGTAGCTGCTGCTGGTGTGGATCATTTCACTG	486	TCCATTACAAAACCTCGATGTGCAGTGTAGCTGCTGCTGGTGTGGATCATTTCACTG	487	TCCATTACAAAACCTCGATGTGCAGTGTAGCTGCTGCTGGTGTGGATCATTTCACTG	488	TCCATTACAAAACCTCGATGTGCAGTGTAGCTGCTGCTGGTGTGGATCATTTCACTG
238	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle	257	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle	258	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle	259	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle
487	GTAGCTGTCTTCGATGACCTTCTTGATCATCATCAACCAAC-----AGGACCAACAGA	540	GTAGCTGTCTTCGATGACCTTCTTGATCATCATCAACCAAC-----AGGACCAACAGA	541	GTAGCTGTCTTCGATGACCTTCTTGATCATCATCAACCAAC-----AGGACCAACAGA	542	GTAGCTGTCTTCGATGACCTTCTTGATCATCATCAACCAAC-----AGGACCAACAGA
258	ThrThrCysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr	277	ThrThrCysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr	278	ThrThrCysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr	279	ThrThrCysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr
541	TCAGCTGTCTCGACCTCACCAGTTTCGAGTGAACCT--AATACTATTAAAGTGGTCAAC	597	TCAGCTGTCTCGACCTCACCAGTTTCGAGTGAACCT--AATACTATTAAAGTGGTCAAC	598	TCAGCTGTCTCGACCTCACCAGTTTCGAGTGAACCT--AATACTATTAAAGTGGTCAAC	599	TCAGCTGTCTCGACCTCACCAGTTTCGAGTGAACCT--AATACTATTAAAGTGGTCAAC
278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr	297	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr	298	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr	299	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr
598	CTGATT--TTGACTTCGAACACTTTCTTCGCTCCCTTGGTGATAGTGACACTTTGCTAT	654	CTGATT--TTGACTTCGAACACTTTCTTCGCTCCCTTGGTGATAGTGACACTTTGCTAT	655	CTGATT--TTGACTTCGAACACTTTCTTCGCTCCCTTGGTGATAGTGACACTTTGCTAT	656	CTGATT--TTGACTTCGAACACTTTCTTCGCTCCCTTGGTGATAGTGACACTTTGCTAT
298	ThrThrLeuIleHisLys-----LeuLysSerLys	307	ThrThrLeuIleHisLys-----LeuLysSerLys	308	ThrThrLeuIleHisLys-----LeuLysSerLys	309	ThrThrLeuIleHisLys-----LeuLysSerLys
655	ACACAGATTATCCACACTCTGACCATCGGATCGAAATGACAGCTGCCCTTAGACGAAA	714	ACACAGATTATCCACACTCTGACCATCGGATCGAAATGACAGCTGCCCTTAGACGAAA	715	ACACAGATTATCCACACTCTGACCATCGGATCGAAATGACAGCTGCCCTTAGACGAAA	716	ACACAGATTATCCACACTCTGACCATCGGATCGAAATGACAGCTGCCCTTAGACGAAA
308	AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIle	327	AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIle	328	AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIle	329	AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIle
715	GCACGA-----AGGCTAACCATTTGCTACTCTCTTGCATTTTACGTA	756	GCACGA-----AG				

[illegible]

RESULT 40

```

US-09-943-798-3
; Sequence 3, Application US/09943798
; Patent No. US20020065215A1
; GENERAL INFORMATION:
; APPLICANT: Glaxo Group Limited
; TITLE OF INVENTION: Polypeptide
; FILE REFERENCE: QG1021
; CURRENT APPLICATION NUMBER: US/09/943,798
; CURRENT FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-943-798-3

```

Alignment Scores:	
Pred. No.:	2.4e-25
Score:	369.50
Percent Similarity:	49.33%
Best Local Similarity:	29.53%
Query Match:	17.30%
DB:	10
Length:	101
Matches:	87
Conservative:	58
Mismatches:	108
Indels:	41
Gaps:	8

US-09-208-629F-3 (1-407) X US-09-943-798-3 (1-1014)

107	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer	126	
Qy			
Db	CTCAAGATGCATACCTCCCTGTTATTATGGCATATTCTTCGTGGGATTTCCAGGC	150	
Qy	127	AsnIleValThrLeu-----TrpLysLeuSerLeuArg	137
Db			
Qy	151	AATGCGATGATGATATCCACTTACATTTTCAAAATGAGACTTGGAG-----	198
Qy	138	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe	157
Db			
Qy	199	-----AGCAGCACCATCATATTGCTGAACTGCGCTGCACAGATCTGCTGAT	246
Qy	158	CysValThrLeuProPheGlyIleAlaTyrHisIleuAsnGlyAsnAsnTrpValPheGly	177
Db			
Qy	247	CTGACGAGCTCCCTTCCTGATTCACTACTACTGCCAGTGGCGAAATCGATCTTTGGA	306
Qy	178	GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	197
Db			
Qy	307	GATTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATACGAGCATCCCTC	366
Qy	198	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	217
Db			
Qy	367	TTTCCTCACCTGTTTCAGCATCTTCGGCTACTGTGTGATCATTCACCCATGAGCTGCTTT	426
Qy	218	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe	237
Db			
Qy	427	TCCATTACAAAACTCGATGTGCAGTTGTAGCCCTGTCTGTGGTGTGGATCAATTCACATG	486
Qy	238	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisIleuValHisSerGluIle	257
Db			
Qy	487	GTAGCTGTCAITTCGGATGACCTTCTTGATCATCAACCAAC-----AGGACCAACAGA	540
Qy	258	ThrThrCysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr	277
Db			
Qy	541	TCAGCCTGTCTCGACTCCACCGATTTCGATGAATC---AATACTATTAGTTGGTCAAC	597
Qy	278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr	297

```
Db 598 CTGATT---TTGACTGCAACTACTTTCTGCTCCCTTGGTGTAGTGACACTTTTGCTAT 654
Qy 298 ThrThrLeuIleHisLys-----LeuLysSerLys 307
Db 655 ACCACGATTATCCACACTCTGACCCATGGAGTGCNAACCTGACAGCTGCCTTAAGCAGAAA 714
Qy 308 AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIle 327
Db 715 GCAGCA-----AGGCTAACCATTCCTGCTACTCCTTGCAATTTACGTA 756
Qy 328 CysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHisAsn 347
Db 757 TGTITTTTACCCCTTCCATATCTTGAGGGTCATTGCGATCGAATCGCCTGCTTTCAATC 816
Qy 348 Thr-----AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
Db 817 AGTTGTTCCATTGAGAAATCAGATCCATGAGCTTACATCGTTTCTAGACCAATTAGCTGCT 876
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
Db 877 CTGNACACCTTTGGTAACCTGTTACTATATGTGTGTGTGCTAGC 918
```

Search completed: June 29, 2003, 11:46:16
Job time : 170.838 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:52:59 ; Search time 45.9436 Seconds
(without alignments)
2716.756 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXQHPVAGSQDKMKIL.....AMARPLXRRPRDIWEDIHAW 407

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09208629/runat 24062003 101626 12944/app query.fasta_1.1166
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CGN 1 1 47 @runat 24062003 101626 12944 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2121	99.3	1224	2	US-08-742-440A-1
2	1443.5	67.6	1224	2	US-08-742-440A-4
3	1369	64.1	1102	2	US-08-742-440A-5
4	1300.5	60.9	1124	2	US-08-742-440A-2
5	584	27.3	2732	1	US-08-476-000-60
6	584	27.3	2732	2	US-08-476-000-60
7	584	27.3	2732	2	US-08-476-976-60
8	584	27.3	2732	4	US-08-474-410-60
9	584	27.3	2732	4	US-08-486-673B-60
10	579.5	27.1	1475	1	US-08-097-938-1
11	579.5	27.1	1475	1	US-08-476-000-1
12	579.5	27.1	1475	1	US-08-472-840-1

13	579.5	27.1	1475	2	US-08-476-976-1	Sequence 1, Appli
14	579.5	27.1	1475	3	US-08-474-410-1	Sequence 1, Appli
15	579.5	27.1	1475	4	US-08-486-673B-1	Sequence 1, Appli
16	563	26.4	1255	1	US-08-097-938-3	Sequence 3, Appli
17	563	26.4	1255	1	US-08-476-000-3	Sequence 3, Appli
18	563	26.4	1255	1	US-08-472-840-3	Sequence 3, Appli
19	563	26.4	1255	2	US-08-476-976-3	Sequence 3, Appli
20	563	26.4	1255	3	US-08-474-410-3	Sequence 3, Appli
21	563	26.4	1255	4	US-08-486-673B-3	Sequence 3, Appli
22	551	25.8	1414	1	US-08-476-000-62	Sequence 62, Appli
23	551	25.8	1414	2	US-08-472-840-62	Sequence 62, Appli
24	551	25.8	1414	2	US-08-476-976-62	Sequence 62, Appli
25	551	25.8	1414	3	US-08-474-410-62	Sequence 62, Appli
26	551	25.8	1414	4	US-08-486-673B-62	Sequence 62, Appli
27	511.5	23.9	1764	1	US-08-313-553-12	Sequence 12, Appli
28	511.5	23.9	1764	3	US-08-767-993-12	Sequence 12, Appli
29	509	23.8	3480	1	US-07-657-769B-68	Sequence 68, Appli
30	509	23.8	3480	1	US-07-789-184-219	Sequence 219, App
31	509	23.8	3480	1	US-08-475-263-219	Sequence 219, App
32	509	23.8	3480	1	US-08-485-886-219	Sequence 219, App
33	509	23.8	3480	2	US-08-477-362-219	Sequence 219, App
34	509	23.8	3480	2	US-08-477-134-219	Sequence 219, App
35	509	23.8	3480	3	US-08-473-489A-219	Sequence 219, App
36	509	23.8	3480	3	US-08-485-695-219	Sequence 219, App
37	509	23.8	3480	4	US-08-018-760-219	Sequence 219, App
38	507.5	23.8	4895	3	US-09-053-866-1	Sequence 1, Appli
39	507.5	23.8	4895	4	US-09-479-130-1	Sequence 1, Appli
40	431.5	20.2	1643	1	US-08-383-750-3	Sequence 3, Appli
41	431.5	20.2	1643	3	US-08-352-678-3	Sequence 3, Appli
42	431.5	20.2	1643	5	PCT-US93-09636-3	Sequence 3, Appli
43	398.5	18.7	1901	1	US-08-153-848-43	Sequence 43, Appli
44	398.5	18.7	1901	3	US-09-299-843A-43	Sequence 43, Appli
45	398.5	18.7	1901	4	US-09-088-337B-43	Sequence 43, Appli

ALIGNMENTS

RESULT 1
US-08-742-440A-1
; Sequence 1, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400

TELEFAX: 650 327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1224 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-742-440A-1

Alignment Scores:

Pred. No.: 2,43e-211 Length: 1224
 Score: 2121.00 Matches: 404
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 2 Gaps: 0

US-09-208-629F-3 (1-407) x US-08-742-440A-1 (1-1224)

```

Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
Db 3 ACTTTGTATACTTAACCAACATCTCTGTAGCCGGGTCTCAGGACATCAAGATGAAATCCCTT 62
Qy 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
Db 63 ATCTTGGTTCAGCTGGGGTCTGCTTCTGCCAGTCACTGTTGGCCAAAGTGGCAATAAT 122
Qy 41 ValSerAspSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
Db 123 GTTTCAGACAACTCAGCAAGCCAACTTAACATTAAAGAGTTTAAATGGGGTGGCCCAA 182
Qy 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80
Db 183 AATACCTTTGAAGAATTTCCCACTTTCTGACATAGAGGGCTGGACAGAGCCACCAACT 242
Qy 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIle 100
Db 243 ATAAAGCGAGTGTCCGAGGACAGATTTTCACTCTCCAGTGAATAATGTACCATA 302
Qy 101 GlyTyrLeuArgSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
Db 303 GGATACCTGAGAAGTCTCTTAAGTACCAAGTATACCTGCCATCTATATCTGCTGTTT 362
Qy 121 ValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer 140
Db 363 GTGGTTGGTGPAACATCCACATCGTGACCTGTGGAAACTCTCTTAAGGACCAATCC 422
Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 423 ATCAGTCTGGTCATCTTTTCACCAACCTGGCCATCGCAGATCTCTTTTCTGTGCACA 482
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 483 CTGCCATTTAAGATCGCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGTCAATG 542
Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
Db 543 TGCCGGATCACCGGTGGTTTCTACGGCAACATGTACTGCGCTATCTCTGATCCCTCACT 602
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 603 TGCATGGCATCAACCGGTACTTGGCCACGGCTCACCTTTTCATACCAAGAGTCCCTCCC 662
Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 663 AAACGCAGCTTCTCTTCTCATGTGTGGCATAGTGTGGGTCACTGTTTCTTATACATG 722
Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
Db 723 CTGCCCTTTGTTCATCTCTTGAAGCAGGAGTACCACCTCGTCCACTCAGAGATCACCCACCTGC 782

```

```

Qy 261 HisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSer 280
Db 783 CACGATGCTGTCGACGGTGGAGTCCCATCATCTCTCCGATTCTACTACTTGGTCTCC 842
Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
Db 843 TTAGCATTTCTTTGGGTTCCTCATCCGTTTGTGATCATCATCTTCTGTGTACAGACTCTC 902
Qy 301 IleHisLysLeuLysSerLysAspArgIleTyrPheGlyTyrIleLysAlaValLeuLeu 320
Db 903 ATCCCAAACTTAAATCAAGGATCGGATATGGTGGGTACATCAAGCCGCTCTCTCTC 962
Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
Db 963 ATCTTGTGATTTTCACAATTTGCTTGGCCCCCAACATCATACTCGTAATCCACCAT 1022
Qy 341 AlaAsnTyrTyrTyHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCys 360
Db 1023 GCCAATCTACTACCACAATACCAGACAGCTTGTATCTTATGTATCTTATGCTCTGTGC 1082
Qy 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
Db 1083 CTGGGAGGCTGAATAGCTGCTAGATCCATCTCTTACTTGTCTATGTCGAAAGTTGTA 1142
Qy 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
Db 1143 GATCAGCTTAATCTTTAGTCGGCAATGGCAAGACCACTTTAGAGACCAAGGAGATATC 1202
Qy 401 TrpGluAspIleHisAlaTrp 407
Db 1203 TGGGAAGACATACATGCTTGG 1223

```

RESULT 2

```

US-08-742-440A-4
; Sequence 4, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Iehihari, Hiroaki
; APPLICANT: Comolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1224 base pairs

```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-742-440A-4

Alignment Scores:
Pred. No.: 6,84e-141 Length: 1224
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 2 Gaps: 1

US-09-208-629F-3 (1-407) x US-08-742-440A-4 (1-1224)

Qy 11 GlySerGlnAspIleTyrMetLysLeuValAlaGlyLeuLeuPheLeu 30
Db 40 GGGACTCAGGTCAACAAATGAAGCCCTCATCTTTCAGCTGCTGGCCCTCCTGCTTCTG 99
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspSerAlaLysProThr 49
Db 100 TTGCCCACTTTTGTGTCAGATGGATGGAATAATGATACAAACACTTGGCCAAAGCCCAACC 159
Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 160 TTACCCATTAAGACCTTTCGTGGAGCTCCCCCAAATCTTTGAAGAGTTCCCTTTTCT 219
Qy 70 AspileGluGlyTyrThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 220 GCCTTGAAGGCTGGAGCAGGACCATTAAGTAAATAATTAAGTCCCTGGAAGAAAGT 279
Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 280 GCTTCACATCTCCATGTGAATAATGCTACCATGGGGTACTGACCACCTCTTAAGTACT 339
Qy 110 GlnValIleProAlaIleTyrIleLeuPheValValGlyValProSerAsnIleVal 129
Db 340 AAATGATACCTGCATCTACCTCTGCTGTTGTAGTGTGTCGGCCCAATGCTGTG 399
Qy 130 ThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 400 ACCCTGGAGTCTTTCTTCAGGACGAGATCCATCTGTACCACCTGTAATCTACACCAAC 459
Qy 150 LeuAlaIleAlaAspLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
Db 460 CTGGCCATTCGAGATTTCTTTTGTGTACATGTCCTTTAAGATAGTTATCATCTC 519
Qy 170 AsnGlyAsnThrValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
Db 520 AATGGGAACAATGGGTATTGGAGAGGTCTGTGCGGGCCACACAGTCATCTTCTAT 579
Qy 190 GlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
Db 580 GGCAACATGATGCTCCTCATCTCTCTGCTGCTGATCAGCATCAACCGCTACCTGGCC 639
Qy 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuMetCys 229
Db 640 ATCGTCCATCTTTACCTACCGGGGCTGCCCAAGCACACCTATGCTTGGTAACTAT 699
Qy 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
Db 700 GGACTGGTGGGCAACAGTTTTCTTATATATGCTGCCATTTTTCATCTAGTGAAGCAGAA 759
Qy 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
Db 760 TATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCACAACTTGGAGTCC 819
Qy 270 ProSerPheArgPheTyrTyrPheValSerLeuAlaPheGlyPheLeuIlePro 289
Db 820 TCATCTCCCTTCCAACTCTATTACTCTCTTGGCATCTTTTGAATCTTAAATTCGA 879
Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAspArg 309
```

```
880 TTTGTCTTATCATCTACTGTCATGCGGACCATCATCGGACACTTAATGCATACGATCAT 939
310 IleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCysPhe 329
940 AGATGGTGTGGTATGTTAAGGCGAGTCTCCTCATCTTGTGATTTTACCAITTTGCTTT 999
330 AlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
1000 GCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAAACAACACTGAT 1059
350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
1060 GGCTTATATTTATATATCTCATAGCTTTGTGCTGGTAGTCTTAATAGTTGCTTAGAT 1119
370 ProPheLeuTyrPheValMetSerLys 378
1120 CCATTCCTTATTCTTCTCATGCAAAA 1146

RESULT 3
US-08-742-440A-5
; Sequence 5, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1102 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-742-440A-5

Alignment Scores:
Pred. No.: 3.28e-133 Length: 1102
Score: 1369.00 Matches: 245
Percent Similarity: 84.96% Conservative: 43
Best Local Similarity: 72.27% Mismatches: 51
Query Match: 64.09% Indels: 0
DB: 2 Gaps: 0
```

```
US-09-208-629F-3 (1-407) x US-08-742-440A-5 (1-1102)
Qy 40 AsnValSerAspSerAlaLysProThrIleLysSerPheAsnGlyGlyPro 59
Db 13 AATGATACAACTTGGGAAGCCAACTTACCCTTAAGACCTTTCGTGAGCTCC 72
Qy 60 GlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 79
Db 73 CCAATTCCTTTGAAGAGTTCCTTCTGCTTGAAGCTGGACAGGACGATT 132
Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThr 99
Db 133 ACTGTAAAAATTAGTCCCTGGAAGAGTCTCCATCTCCATGTGAAAAATGCTACC 192
Qy 100 IleGlyTrpLeuArgSerSerLeuSerThrGlnValIleProAlaIleThrIleLeu 119
Db 193 ATGGGGTACCTGACCACTCTTAAGTACTAAATGATGATCTACCTACCTACCTGGTG 252
Qy 120 PheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys 139
Db 253 TTGTAGTGTGTCCTGGGCAATGCTGTGACCTGTGGATGCTTTCTTCAGAGACGAGA 312
Qy 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
Db 313 TCCATCTGTACCACTGTATTCTACCACTGSCCATTCAGATTTCCTTTTGTGTT 372
Qy 160 ThrLeuProPheLysIleAlaLysHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
Db 373 ACATGTGCTTTTAAAGTAGCTTATCATCTCAATGGGAACAACCTGGGTATTGGAGAGGC 432
Qy 180 MetCysArgIleThrValValPheThrGlyValMetThrGlyValIleLeuLeu 199
Db 433 CTGTGCGGGCCACCACTGCTTCTATGGCAACATGACTGCTCCATCTGCTCCTT 492
Qy 200 ThrCysMetGlyIleAsnArgTrpLeuAlaThrAlaHisProPheThrTrpGlnLysLeu 219
Db 493 GCCTGCATCAGCATCAACCGCTACTGCGCATCTGCTCCATCTTTCACCTACCGGGCGTG 552
Qy 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTrp 239
Db 553 CCCAAGCACACTTATGCTTGGTAACATGGAGTGGTGGGCAACAGTTTTCTTATAT 612
Qy 240 MetLeuProPheValIleLeuLysGlnGlyHisLeuValHisSerGluIleThrThr 259
Db 613 ATGCTGCCATTTTCATCTAGCAGCAGGAATATATCTTGTTCAGCCAGACATCACCACC 672
Qy 260 CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTrpPheVal 279
Db 673 TGCCATGATGTTCAACAACACTTGGAGTCCCTCATCTCCCTTCCAACTCTATTACTTCATC 732
Qy 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTrpThrThr 299
Db 733 TCCTTGGCATCTTTTGGATTCTTAATTCATTTGTGCTTATCATCTACTGCTATGACGCC 792
Qy 300 LeuIleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTrpIleLysAlaValLeu 319
Db 793 ATCATCGGACACTTAATGATCATGATCATAGATGGTTGGTATGTTAGGCGAGTCTC 852
Qy 320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHis 339
Db 853 CTCATCTTGTGATTTTACCATTGCTTGTCTCCAAGCAATATATTCTTATTATTATTCAC 912
Qy 340 HisAlaAsnTrpTrpTrpHisAsnThrAspSerLeuTrpPheMetTrpLeuIleAlaLeu 359
Db 913 CATGCTAACTACTACTACCAACACTGGTGGCTTATATTATATATCATCATAGCTTTG 972
Qy 360 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTrpPheValMetSerLys 378
Db 973 TGCTGGGTAGTCTTAATAGTTGCTTAGATCCATCTCTTATTATTCTCATGTCAAAA 1029
US-08-742-440A-2
RESULT 4
```

```
Sequence 2, Application US/08742440A
Patent No. 5892014
GENERAL INFORMATION:
APPLICANT: Coughlin, Shaun
APPLICANT: Ishihari, Hiroaki
TITLE OF INVENTION: Protease Activated Receptor
TITLE OF INVENTION: 3 and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,440A
FILING DATE: 30-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: UCAL/060PAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1124 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
US-08-742-440A-2
Alignment Scores:
Pred. No.: 4,48e-126 Length: 1124
Score: 1300.50 Matches: 265
Percent Similarity: 71.99% Conservative: 10
Best Local Similarity: 69.37% Mismatches: 16
Query Match: 60.88% Indels: 92
DB: 2 Gaps: 4
US-09-208-629F-3 (1-407) x US-08-742-440A-2 (1-1124)
Qy 29 PheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAlaLysPro 48
Db 15 TTCCTTTCAATACAC-----GGCATAAATGTTTCAGACAACTCAGCAAGCCCA 62
Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 63 ACCTTAACATTATAAGAGTTTTAATGGGGTCCCAAAATACCTTTTCAAGAAATTCNN--- 118
Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 118 ----- 118
Qy 89 SerIleSerThrLeuHisValAsnAlaThrIleGlyTrpLeuArgSerSerLeuSer 108
Db 119 NNNNTACAACTCTCCATGTGTAATGCTACCTGGGATACCTGAGAGTTCCTTAAGT 178
Qy 109 ThrGlnValIleProAlaIleTrpIleLeuPheValGlyValProSerAsnIle 128
```



```
Db 436 GCCGACCTCTCTCTGTCATCTGGTCCCTCGAAGATCTCCACCACTACATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgileThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy 193 TyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Db 556 TATTGCTCCATCTCTTTCATGACCTGCTCAGCGTCAGAGGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrTyrGlnTyrLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCATGGGACACCCAGG---AAGAAGGCAAAACATCCCGTGGCGTCTCTTGGCAATC 672
Qy 233 TrpValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeu 252
Db 673 TGGCTCTGATTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluLeuThrThrCysHisAspValValAspAlaCysGluSerProSer 272
Db 733 CCAGCATTTGAACATCACCATCTGTCCAGCATGTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTCTAAATTACTTCTCTACTGCGCAATTGGAGTCTTCTGTTCCCGGCCCTCCT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTCGCTACGTCTCATGATCAAGCGTCCGCTCTTCTGATGGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeu 322
Db 910 TCAGAGAAGAAAGGAGAGGGCT-----ATCCGACTCATCATCACCGTCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGTACTTCTCTGCTTCTCTAGCAACCTCTCTGCTGATGTCATTATTTCCTA 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCCAGGAGAGCCAGCTGTACGCCCTCTACCTTGTGCGCCCTCTGCTCTG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGTCATAGACCCCTTTGCTCTATTACTTTGCTCAAAA 1125
```

RESULT 7

```
US-08-476-976-60
; Sequence 60, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..1269
; US-08-476-976-60

Alignment Scores:
Pred. No.: 5,21e-51 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 2 Gaps: 10

US-09-208-629f-3 (1-407) x US-08-476-976-60 (1-2732)
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db 61 CCAGTCCCGGGGTGGAGAGTCTCAGCCTGGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db 112 ACCCTTCTGGCGCCTCGGTCTCTCTCAGCGGACCGAGAACCTTGCCCGGACGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrLeuLys 53
Db 172 AACAGTAAAGGAAGAGTCTTATTGCGAGATTAGAACCCAGCCTCAATCATCTGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACCGCTTTTCCATCGATGAGTCTCTCGCTCCATCTCACCGGG 291
Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db 292 -----AAGCTGACCACGGTCTTTCTTCG 315
Qy 114 AlaIleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTyrLys 133
Db 316 GTCTCTACATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCCTTTTCCGAACGAAGAAACACCCCGCGTGTATTATACATGGCAACCTGGCCTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db 436 GCCGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy 193 TyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
```

556 TATTGCTCCATCTCTTCATGACCTGCTCAGCGTGCGAGAGGTACTGGGTGATCGTGAAC 615
213 ProPheThrTyrGlnIleValProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
616 CCCATGGGACACCCAGG---AAGAAGGCAAAACATCGCGTTCCTGGCGTCTCTTGGCAATC 672
233 TrpValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeu 252
673 TGGCTCTGATTTTCTGGTCCACCATCTTTGTATGTCATGAAGCAGACCATCTACATT 732
253 ValHisSerGluIleThrCysHisAspValValAspAlaCysGluSerProSerSer 272
733 CCAGCATTTGAACATCACCATCTGTCACGATGTGCTGCTGAG---GAGGTATTGGTGGG 789
273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
790 GACATGTTCAATTTCTCTCTACTGCGCATTTGGAGTCTCTGTTCCCGGCCCTCTT 849
293 IleIlePheCysTyrThrThrLeuIle-----His 302
850 ACTGCATCTGCTGCTGCTCATGATCAAGACGCTCGCTCTCTCTGCTATGGATGAACAC 909
303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
910 TCAGAGAAGAAAGGACAGGGCT-----ATCCGACTCATCATCACCGTGTG 957
323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsn 342
958 GCCATGACTTCACTCTCTGCTCTCTAGCAACCTCTGCTGCTAGTGCATTATTTCTTA 1017
343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
1018 ATCAAAACCAGGAGGAGCCGCTGCTAGCCCTCTACCTTGTGCGCCCTCTGCTGTG 1077
363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
1078 ACCCTCAACAGCTCATAGACCCCTTTGTTCTATTACTTTGTTCTCAAAA 1125

RESULT 8

US-08-474-410-60
Sequence 60, Application US/08474410
Patent No. 6043212
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 2732 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
NAME/KEY: CDS
LOCATION: 73..1269
US-08-474-410-60
Alignment Scores:
Pred. No.: 5,21e-51 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 3 Gaps: 10
US-09-208-629f-3 (1-407) x US-08-474-410-60 (1-2732)
QY 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
DB 61 CCACGTCGCGGGATCGAAGTCTCAGCGTGGCGTGGCTGCTG-----GGAGGTATC 111
QY 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
DB 112 ACCCTTCTGGCGGCTCGTCTCTGCGACCGGACCGAAGACCTTCGACCGGACGCAAC 171
QY 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
DB 172 AACAGTAAGGAGAAAGTCTTATTGGCAGATTAGAAACCCAGCTCCAATCATCTGGGAAA 231
QY 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGly 73
DB 232 GGGTTCCGGTAGAACCGAGCTTTTCCATCGATGAGTTCTCTGCGTCCATCTCACC 291
QY 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
DB 291 ----- 291
QY 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
DB 292 -----AAGCTGACACAGGTCTTCTTC 315
QY 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
DB 316 GTCGCTACATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
QY 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
DB 376 TTCTTTTCCGAACGAAGAAACACCCCGCGCTGATTATCATGTGCCAACCTGGCTTG 435
QY 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
DB 436 GCGGACCTCTCTCTCTCATCTGGTTCCCGCTGAGATCTCTTACCACCTACATGGAAC 495
QY 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
DB 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTCTCATTTGGCTTTTCTATGTTAATG 555
QY 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
DB 556 TATTGCTCATCTCTTCTCATGACCTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 615
QY 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
DB 616 CCATGGGACACCCAGG---AAGAAGGCAAAACATCGCGTGGCGTCTCTTGGCAATC 672
QY 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252

```
Db 673 TGGCTCTGATTTTCGGTCACATCCCTTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATCACCACCTGTCAGATGTCGCTGAG--GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValle 292
Db 790 GACATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTCCCTACGCTCATCATGATCAAGACGCTCGCTCTCTCTCTCTCTCTCTCT 909
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrLeuGlyTyrLeuValLeuLeuLeu 322
Db 910 TCAGAGNAGAAAGGCAGAGGCT-----ATCCGATCATCATCAGCGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCAGAGGAGCCAGCTACGCTCTACCTTGTGCGCTCTCTCTCTCTCTCTCT 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTAATTAATTAATTAATTAATTAAT 1125
```

RESULT 9

```
US-08-486-673B-60
; Sequence 60, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486.673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 60
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)..(1269)
; OTHER INFORMATION: C140 receptor, cDNA and deduced protein sequences
US-08-486-673B-60
```

Alignment Scores:

```
Pred. No.: 5,21e-51 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 4 Gaps: 10
```

US-09-208-629F-3 (1-407) x US-08-486-673B-60 (1-2732)

```
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeu 27
Db 61 CCAGTCCGGGATGCCAAGTCTCAGCTCGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
```

```
Db 112 ACCCTTCTGGCGCCTCGCTCTCTCTGACCGCAGAACCTTCACCGGACGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 ACAGTAAAGGAAGAGTCTTATTGGCAGATTAGAAACCCAGGCTCAATCACTCGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACAGGCTTTTCCATCGATGAGTCTCTGCGTCCATCTCTCACCGG 291
Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db 292 -----AAGTGACCAAGGCTCTTCTTCCG 315
Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
Db 316 GTCGCTACATTAATGTTGTTGATGTTGTTGCCAGTAATGGCATGGCCCTCTGATC 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCCTTTTCCGAACGAAGAAACACCCCGCGTGATTATACATGGCCCAACCTGGGCTTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db 436 GCGGACCTCTCTCTGTCATCTGGTTCCTCCCTCAAGATCTCTACACCATCATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCGCTGTCAAGGTCTCATTTGGCTTTTCTATGGTAACATG 555
Qy 193 TyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrIleuAlaThrAlaHis 212
Db 556 TATTGCTCCATCTCTTTCATGACCTCCCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCCATGGGACACCCAGG---AAGAAGGCAACATCGCGTTGGCGTCTCTCTGGCNAATC 672
Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
Db 673 TGGCTCCTGATTTTCTGGTCAACATCCCTTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATCACCACCTGTCACGATGTCTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTCCCTACGCTCATCATGATCAAGACGCTCGCTCTCTCTCTCTCTCTCTCT 909
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeuLeu 322
Db 910 TCAGAGNAGAAAGGCAGAGGCT-----ATCCGATCATCATCAGCGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCAGAGGAGCCAGCTACGCTCTACCTTGTGCGCTCTCTCTCTCTCTCTCT 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
```

Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1125

RESULT 10

US-08-097-938-1

; Sequence 1, Application US/08097938

; Patent No. 5629174

; GENERAL INFORMATION:

; APPLICANT: SUNDELIN, JOHAN

; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS

; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/097,938

; FILING DATE: 26-JUL-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: MURASHIGE, KATE H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 22803-20006.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1475 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 232..1416

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 232

; US-08-097-938-1

Alignment Scores:

Pred. No.: 6,25e-51

Score: 1475

Matches: 111

Percent Similarity: 60.00%

Best Local Similarity: 37.63%

Query Match: 27.13%

DB: 1

Gaps: 5

US-09-208-629f-3 (1-407) x US-08-097-938-1 (1-1475)

Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114

Db 406 ATCGATGAGTTCTCGCTCCATCTCCACCGGAAGCTGACCGGTCTTCTTCGGTC 465

Qy 115 IleTyrIleLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeu 134

Db 466 GTCTACATATTGTGTTGTGATTGTTGCCCAAGTAAATGCGATGCCCTCGATCTC 525

Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153

Db 526 CTTTTCGACGACGAGAGAAACACCCCGCGTGAATTACATGGCCACCTGGCCCTGGCC 585

Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173

Db 173

Db 586 GACCTCTCTCTGTCTCATCTGGTTCCTCCCTGGAAGATCTCTACCACCTACATGCGCAACAAC 645

Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193

Db 646 TGGGCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGAT 705

Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213

Db 706 TGCTCCATCTCTTCATGACCTGCTCAGCGTGACAGGTACTGGGTGATCGTGAACCCC 765

Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuMetCysGlyIleValTrp 233

Db 766 ATGGGACACCCAGG---AAGAAGGCAACATCGCGTTGGCGTCTCTCTGGCAATCTGG 822

Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnIleThrHisLeuVal 253

Db 823 CTCCTGATTTTCTGCTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATCCA 882

Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273

Db 883 GCATTGAACATCACCACCTGTTCAGATGTGCTGCCTGAG---GAGGTATTGGTGGGGAC 939

Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293

Db 940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTCTCTGTTCCCGCCCTCTTACT 999

Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306

Db 1000 GCATCTGCTCAGTGTCTCAATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059

Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323

Db 1060 GAGAACAAAGGACGAGGCT-----ATCCGACTCATCATCCGCTGCTGCC 1107

Qy 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343

Db 1108 ATGTACTTCACTCTGCTTCTCTAGCAACCTTCTGCTGATGATTTTCTTAATC 1167

Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuAlaLeuCysLeuGlySer 363

Db 1168 AAAACCCAGGACGAGCCACGCTACGCCCTCTACCTTGTGCGCCCTGCTGCTGAC 1227

Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378

Db 1228 CTCACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 11

US-08-476-000-1

; Sequence 1, Application US/08476000

; Patent No. 5716789

; GENERAL INFORMATION:

; APPLICANT: SUNDELIN, JOHAN

; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,000

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 232..1416
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 232
;
US-08-476-000-1
Alignment Scores:
Pred. No.: 6,25e-51 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: Gaps: 5
US-09-208-629F-3 (1-407) x US-08-476-000-1 (1-1475)
QY 95 ValAenAenAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
DB 406 ATCGATGAGTCTTCGCGTCATCTACCGGAGCTGACCGGCTTCCTTCCTCGGTC 465
QY 115 IleTyrIleLeuLeuPheValValcylValProSerAsnIleValThrLeuTyrLysLeu 134
DB 466 GTCTACATATTGTTGTTGATGGTTGGCCAGTAATGGCATGGCCCTCTGGATCTTC 525
QY 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAenLeuAlaIleAla 153
DB 526 CTTTTCGGAAGCAAGAAACACCCCGCTGATTACATGGCCACCTGGCTTGGCC 585
QY 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAenGlyAenAsn 173
DB 586 GACCTCCTCTGTCTGTCCTGTCCTCCCTGAAGATCTCTACCACTTACATGGCAACAAC 645
QY 174 TrpValPheGlyGluValMetCysArgIleThrValValPheTyrGlyAsnMetTyr 193
DB 646 TGGGTCTAGGGGAGGCCCTGTGAAGTGTCTATTGGCTTTTCTATGTTAATCATGTAT 705
QY 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
DB 706 TGCTCCATCTCTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGGTATCGTGAACCC 765
QY 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
DB 766 ATGGGACACCCCAAG---AAGAAGCAACATACCTGCGCTCTCTTGGCAATCTGG 822
QY 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
DB 823 CTCCTGATTTTCTGGTCACCATCTTGTATGTCATGAAGCAGACCATCTACATTCGA 882
QY 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
DB 883 GCATTGAACATCACCACTGTGACGATGTGCTGCTGAG---GAGGTATTGGTGGGGAC 939
QY 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
DB 940 ATGTTCAATTAATCTCTCTCACTGCAATGGAGTCTCTGTTCCCGCCCTCTTACT 999

```


Db 1000 GCATCTGCCTACGGTCTCATGATCAAGACGCTCGGCTCTTCTGCTATGGATGAACACTCA 1059
 Qy 307 -----LysAspArgIleThrLeuGlyTyrIleLysAlaValLeuLeuVal 323
 Db 1060 GAGAACAAAAGGAGGAGGCT-----ATCGACTCATCATCAGCGTGGTGGCC 1107
 Qy 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
 Db 1108 ATGTACTTCACTGCTTTGCTCTAGCAACTTCTGCTGCTAGTGATATTTCCTAATC 1167
 Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySer 363
 Db 1168 AAAACCCAGAGGAGGAGGAGGCTGCTACGCTCTTACCTTGTGCGCTCTGCTGCGACC 1227
 Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerIys 378
 Db 1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 15

US-08-486-673B-1
 ; Sequence 1, Application US/08486673B
 ; Patent No. 6297026
 ; GENERAL INFORMATION:
 ; APPLICANT: Sundelin, Johan
 ; APPLICANT: Scarborough, Robert M.
 ; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
 ; FILE REFERENCE: 44481-5006-08-US
 ; CURRENT APPLICATION NUMBER: US/08/486,673B
 ; CURRENT FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: US 08/097,938
 ; PRIOR FILING DATE: 1993-07-26
 ; PRIOR APPLICATION NUMBER: PCT/US94/08536
 ; PRIOR FILING DATE: 1994-07-26
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1475
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (232)..(1416)
 ; OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein
 ; OTHER INFORMATION: sequences
 US-08-486-673B-1

Alignment Scores:

Pred. No.: 6,25e-51 Length: 1475
 Score: 579.50 Matches: 111
 Percent Similarity: 60.00% Conservative: 66
 Best Local Similarity: 37.63% Mismatches: 101
 Query Match: 27.13% Indels: 17
 DB: 4 Gaps: 5

US-09-208-629F-3 (1-407) x US-08-486-673B-1 (1-1475)

Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
 Db 406 ATCGATGAGTTCTCGCTCCATCTCCCGGGAAGCTGACCACGGTCTTTCTTCGGTC 465
 Qy 115 IleTyrIleLeuLeuValGlyValProSerAsnIleValThrLeuTyrPheLeu 134
 Db 466 GTCTACATATTGTGTGTGATTGTTGCTCCAGTAATGCGATGGCCCTCTGATCTTC 525
 Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
 Db 526 CTTTTCGGAACGAAGAAGAACACCCCGCGTGTATTACATGGCCACCTGGCTTGGCC 585
 Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
 Db 586 GACCTCTCTGTCTGTCTGTTCTCCCTTGAAGATCTCTACCACTACATGCAACAAAC 645

Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
 Db 646 TGGGTCTACGGGAGGCGCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGAT 705
 Qy 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
 Db 706 TGCTCCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 765
 Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
 Db 766 ATGGAGACCCCGG---AAGAAGCAAAACATCGCGTGGCGTCTCTCTGGCAATCTGG 822
 Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
 Db 823 CTCCTGATTTTCTGTGTCACCATCCCTTTGTATGTCATGAGCAGACCATCTACATCCA 882
 Qy 254 HisSerGluIleThrThrCysHisAspValValAlaCysGluSerProSerSerPhe 273
 Db 883 GCATTGAACATCACCACCTGTCCAGATGTCTGCTCGCTGAG---GAGGTATTGGTGGGGAC 939
 Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293
 Db 940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTCCCGCCCTCTCTTACT 999
 Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
 Db 1000 GCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059
 Qy 307 -----LysAspArgIleThrLeuGlyTyrIleLysAlaValLeuLeuVal 323
 Db 1060 GAGAACAAAAGGAGGAGGCT-----ATCGACTCATCATCAGCGTGGTGGCC 1107
 Qy 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
 Db 1108 ATGTACTTCACTGCTTTGCTCTAGCAACTTCTGCTGCTAGTGATATTTCCTAATC 1167
 Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
 Db 1168 AAAACCCAGAGGAGGAGGAGGCTGCTACGCTCTACCTTGTGCGCTCTGCTGCGACC 1227
 Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerIys 378
 Db 1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 16

US-08-097-938-3
 ; Sequence 3, Application US/08097938
 ; Patent No. 5629174
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
 ; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/097,938
 ; FILING DATE: 26-JUL-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22803-20006.00

US-09-208-629F-3 (1-407) x US-08-472-840-3 (1-1255)

QY 49 ThrLeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGluPheProLeu 68
DB 116 ACCCTTGCTTCCTTCTGTACAGAAACCAATAGATCCTCTAAAGGAAGACGCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
DB 176 GGTAAGGTTGATGCACATCCACGTCACGTAAGAGGAGTTACAGTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 227 -----ACAGTCTTTCGTGGATGAGTTTCTGCATCTGCTCACTGGA 271
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValPro 125
DB 272 AAACGTGACCATGCTTCCTCCCAATGCTACACAATGTTGTTGGTGGGTTGGCA 331
QY 126 SerAsnIleValThrLeuThrLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
DB 332 AGTAACGGCATGGCCCTGTGGTCTTCTTTCCGAACCTAAGAAAGACACCCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 392 ATTACATGGCCAACTGGCTGGCTGACCTCTCTCTCACTGGTCCCTTGAAG 451
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 452 ATGGCTTATCACATACATGCAACCACTGGATTTATGGGAAGCTCTTGTATGTCCT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 512 ATGGCTTTTCTATGGCAACATGACTGTTCCATCTCTCATGACCTGCCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 572 CAGAGGTATGGGTATCGTGAACCCCATGGGCGCATCCAGG---AAGAAGGCAACATT 628
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 629 GCCATTGGCATCCTCCCTGGCAATATGCTGCTGATTCTGTGTCACCATCCCTTTGTAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
DB 689 GTCGTGAAGCAGACCATCTTCTCTGCTGCTGCAACATCAGACCTGTCATGATGTTTG 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
DB 749 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
DB 794 CTGGCCATTGGGTCTTCTGTTCCTCCAGCCTTCTCTACAGCCTCTGCTATGCTGATG 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
DB 854 ATCAGAATGCTGGCATCTTCTGCCATGGATGAAACTCAGAGAAGAAAGAGAGGGCC 913
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
DB 914 ATCAAACTCATGTGCTGCTGGCCATGTACTGATCTGCTTCACTCTCCTAGTAACCTT 973
QY 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
DB 974 CTGCTTGGTGCAATTATTCTGATTAAAGCCAGGCGCCAGACCATGCTATGCGCTG 1033
QY 355 TyrIleAlaLeuCysLeuLysSerLeuAsnSerCysLeuAspPropheLeuTyrPhe 374
DB 1034 TACATTGAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTGTCTATTAC 1093
QY 375 ValMetSer 377
DB 1094 TTTGTTTCA 1102

RESULT 19

US-08-476-976-3
; Sequence 3, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..1249
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 56
; US-08-476-976-3
Alignment Scores:
Pred. No.: 2,57e-49 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 7 Gaps:
US-09-208-629F-3 (1-407) x US-08-476-976-3 (1-1255)
QY 49 ThrLeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGluPheProLeu 68
DB 116 ACCCTTGCTTCCTTCTGTACAGAAACCAATAGATCCTCTAAAGGAAGACGCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
DB 176 GGTAAGGTTGATGGCAGATCCACGTCACGTAAGAGGAGTTACAGTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 227 -----ACAGTCTTTCGTGGATGAGTTTCTGCATCTGCTCACTGGA 271

```
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 272 AAACGACCACTGCTCCCTCCCAATGTCTACAAATGTGTGGTGGGTGGTCCCA 331
QY 126 SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
DB 332 AGTAACGGCATGGCCCTGTGGTCTTCTTCGAACTAAGAAGAACGACCCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 392 ATTTACATGCCCACTCTGGCTGGCTGACCTCTCTCTCATCTGGTTCCTCCCTTGAAG 451
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr 184
DB 452 ATTGCCCTATACATATGCACTGCAACCACTGGATTTATGGGAAGCTCTTTGTATGTGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIle 204
DB 512 ATTGGCTTTTCTATGCAACATGACTGTTCCTCATCTTCATGACCTGCCCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 572 CAGAGGTATTGGGTATCGTCAACCCCATGGGCACCTCCAGG---AAGAAGGCAACACT 628
QY 225 SerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuProPheVal 244
DB 629 GCCATTGGCATCTCCCGCAATATGCTGCTGATCTCTGCTGACCATCCCTCTGTAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 689 GTCGTGAAGCAGACCATCTTCACTTCTGCTGCTGCAACATCAGACCTGCTCATGTTTG 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
DB 749 -----CCTGAGCAGCTCTTGTGGGAGACATGTTCAATTACTTCTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
DB 794 CTGGCCATTGGGTCTTCTGTTCCAGCCTTCTCACAGCCTCTGCTATGCTGTGAT 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTyrLeuGlyTyr 314
DB 854 ATCAGATGCTGCGATCTCTGCGATGATGAACTCAGAGAAGAAAGAGAGAGGCC 913
QY 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
DB 914 ATCAAACTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
QY 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
DB 974 CTGTTGTTGGTGCAATATTTCTGATTAAGCCAGCCAGCCAGCCATGCTATGCCCTG 1033
QY 355 TyrLeuIleAlaCysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPhe 374
DB 1034 TACATTGAGCCCTCTGCTCTCTACCTTAAACAGCTGCATCGACCCCTTTGCTATTAC 1093
QY 375 ValMetSer 377
DB 1094 TTTGTTTCA 1102
RESULT 20
US-08-474-410-3
; Sequence 3, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
```

```
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 56..1249
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 56
US-08-474-410-3
Alignment Scores:
Pred. No.: 2,57e-49 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 7 Gaps: 7
US-09-208-629F-3 (1-407) x US-08-474-410-3 (1-1255)
```

```
QY 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
DB 116 ACCCTTGCTCTCTCTTCTTGTACAGGACCAATAGATCTCTTAAAGGAAGACGCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
DB 176 GGTAAGGTTTGTGGCATCCACGTCCTGGGAAAGGAGTTACAGTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
DB 227 -----ACAGTCTTTTCTGTGGATGAGTTTCTGTCATCTGTCTCCTCCTGGA 271
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 272 AAACGACCACTGCTCTCTCTTCCATTTGCTACAAATGTGTGGTGGGTGGTCCCA 331
QY 126 SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
DB 332 AGTAACGGCATGGCCCTGTGGTCTTCTTCGAACTAAGAAGAACGACCCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 392 ATTTACATGCCCAATCTGGCTGGCTGACCTCTCTGTCATCTGTGTTCCCTTGAAG 451
```

QY 165 IleAlaTyrHisLeuAenGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGGCTATCACATACATGCAACAACTGATTTATGGGAAGCTCTTTGTAATGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 512 ATTGGCTTTTCTATGCAACATGCTGTCATTTCTTATGACCTCCCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 572 CAGAGGTATTGGGTTCATCGTGAACCCATCGGGCACTCCAGG--AAGAAGGCAAAACATT 628
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 629 GCCATTGGCATCTCCCTGGCAATATGCTGCTGATTTCTTGTGTCACCATCCCTTTGTAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 689 GTGCTGAAGCAGACCATCTTCATCTCCCTGACCATCAGACCTGTCATGATGTTTG 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 749 -----CCTGAGCAGCTCTTGTGGGAGACATGTTCAATTACTTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleValTrpValMetValPheLeuTyrMetLeuProPheVal 300
Db 794 CTGGCAATGGGGTCTTCTGTTCCAGCCTTCTCCACAGCCTCTGCTGCTATGCTGATG 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 854 ATCAGAAATGCTGGCATCTTCTGCAATGATGAAACTCAGAGAAGAAAGAGAGGGCC 913
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 914 ATCAAACTATTGCTACATGCTGCTGGCCATGATGATCTGCTTCACTCTAGTAACCTT 973
QY 335 IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 974 CTGCTTGTGGTCAATTTCTGATTAAAGAGCCAGGCGGAGCCATGCTATGCCCTG 1033
QY 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
Db 1034 TACATTGAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTTGCTATTAC 1093
QY 375 ValMetSer 377
Db 1094 TTTGTTTCA 1102

RESULT 21

US-08-486-673B-3
; Sequence 3, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(1249)
; OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein

; OTHER INFORMATION: sequences
US-08-486-673B-3

Alignment Scores:
Pred. No.: 2,57e-49 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 4 Gaps: 7

US-09-208-629f-3 (1-407) x US-08-486-673B-3 (1-1255)

QY 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 116 ACCCTTGTCTTCTTCTTGTACAGGAACCAATAGATCTCTAAAGGAAGACCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
Db 176 GGTAAAGTTGATGGGCACATCCACGTCACCTGGAAGAGGAGTTACAGTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
Db 227 -----ACAGTCITTTCTGGATGAGTCTTCTGCATCTGCTCCTCACTGGA 271
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 272 AAATCACCACCTGCTTCTTCCCAATGCTACACAATGCTGTTGTGGTGGGTGGCCA 331
QY 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
Db 332 AGTAACGGCATGCCCTGTGGGTCTTCTTCTCCGAACAAAGAAAGACACCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 392 ATTACATGGCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGTTGCCCTTGAAG 451
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGCCTATCACATACATACGCAACAACTGATTTATGGGAAGCTCTTTGTAATGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 512 ATTGGCTTTTCTATGGCAACATGATGTTCTTCTTCAATGACCTGCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 572 CAGAGGTATTGGGTTCATCGTGAACCCATCGGGCACTCCAGG--AAGAAGGCAAAACATT 628
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 629 GCCATTGGCATCTCCCTGGCAATATGCTGCTGATTTCTGCTGCTCACCCTTTGTAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 689 GTGCTGAAGCAGACCATCTTCTTCTCCCTGACCATCAGACCTGCTCATGATGTTTG 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 749 -----CCTGAGCAGCTCTTGTGGGAGACATGTTCAATTACTTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleValTrpValIleIlePheCysTyrThrThrLeu 300
Db 794 CTGGCAATGGGGTCTTCTGTTCCAGCCTTCTCAGACCTCTGCTCATGCTGCTGATG 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 854 ATCAGAAATGCTGGCATCTTCTGCAATGATGAAACTCAGAGAAGAAAGAGAGGGCC 913
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 914 ATCAAACTATTGCTACATGCTGCTGGCCATGATGATCTGCTTCACTCTAGTAACCTT 973

APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476.976

FILING DATE: 07-JUN-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/390,301

FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20

TELECOMMUNICATION INFORMATION
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS
LENGTH: 1414 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:
NAME/KEY:
LOCATION:
FEATURE:
NAME/KEY:
LOCATION:
8-476-976-62

Length:	5,41e-48	1414
Matches:	51.0%	118
Conservative:	51.7%	72
Mismatches:	32.1%	139
Indels:	25.8%	38
Gaps:	2	8

9-208-629F-3 (1-407) x US-08-476-976-62 (1-1414)

22 LeuValAlaIadGlyLeuIeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 ||| ||| ||| ||| ||| ||| ||| |||
 74 CTGTGGGGCCCACTCTGTAGACGCCTCTCTCTCGATCGCACCATCCAAGA 133
 ||||| |||||

42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
 ||| ||| ||| ||| ||| ||| ||| |||
 134 ACCAATAGATCTCTAAAGGAAGACCCCTATTGTGTAAGTGTGATGCC----- 181
 ||||| |||||

62 ThrPheGluGluPheProLeuSeraspIleGluGlyTrpThrGlyAlaThrThrIle 81
 ||| ||| ||| ||| ||| ||| ||| |||
 182 -----ACATCCCACGTC-----ACTGAAAAAGAGGTTACAGTT 214
 |||||

82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGly 101
 ||| ||| ||| ||| ||| ||| ||| |||
 215 GAA-----ACAGTCTTTTCTGTGGATGATGTTTTCTTCGCATCT 250
 |||||

102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleThrValLeuLeuPheVal 121
 ||| ||| ||| ||| ||| ||| ||| |||

```

; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..1240
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 50
; US-08-474-410-62

Alignment Scores:
Pred. No.: 5,41e-48 Length: 1414
Score: 551.00 Matches: 118
Percent Similarity: 51.77% Conservative: 72
Best Local Similarity: 32.15% Mismatches: 139
Query Match: 25.80% Indels: 38
DB: 3 Gaps: 8

US-09-208-629F-3 (1-407) x US-08-474-410-62 (1-1414)

QY 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
Db 74 CTGCTGGGGCCGCCATCTCTAGAGCCCTCTCTCTCTGAGTGGCCACCATCCCAAGGA 133
QY 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
Db 134 ACCAATAGATCCTTAAGGAGNAGCCATTATGCTTAAGTTGATGCC-----181
QY 62 ThrPheGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
Db 182 -----ACATCCACGTC-----ACTGAAAGAGAGTTACAGTT 214
QY 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
Db 215 GAA-----ACAGTCTTTCTGTGATGAGTTTCTGCACT 250
QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
Db 251 GTCTCTGCTGGAACCTGACCATGCTCTCTCTCCATTTCTACACAATTGTTGCG 310
QY 122 ValGlyValProSerAsnIleValThrTriPysLeuSerLeuArgThr---LysSer 140
Db 311 GTGGGTTTGCAAGTAACGGCATGGCCCTATGGCTCTTTCTTCCGAACCTAAGAGAAG 370
QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 371 CACCCTCTGTGATTTACATGGGCAATCTGGCCTTGGCTTGGCTCTCTCTCTGTCATCG 430

```

```

161 LeuProPheIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
431 TTCCCTTGAAGATTGCCTATCACATACATGCAACAACCTGGATTTATGGGAAGCTCTT 490
181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
491 TGTAAATGTGCTTATTGGCTTTTCTATCGCAACATGACTCTGTTCCATTCTCTTCATGACC 550
201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
551 TGCCTCAGTGTGCAGAGGTATTGGGTCATCGTGAACCCCATGGGGCACTCCAGG---AAG 607
221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
608 AAGGCAACATTCGCATTCCTCGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667
241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
668 ATCCCTTTGTATGTGCTGAAGCAGACCATCTTCATTCCTGCCCCGCAACATCAGACCTGT 727
261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
728 CATGATGTTTG-----CCTGAGCAGCTCTTGTGGGAGACATGTTCAAT 772
277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
773 TACTTCTCTCTGCGCATTCGGGCTTTTCTGTTCCAGCCTTCTCAGACCTCTGCCC 832
297 TyrThrThrLeuIleHisLysLysSer-----LysAspArgIle 310
833 TATGTCGTGATGATCAGAAATGCTGCGATCTTTCGCATGATGATGATAAACTCAGAGAAGAA 892
311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
893 AGGAGAGGGCCATCAAACTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
953 CCTAGTAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
1013 GTCTATGCCCTGTACATGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCC 1072
371 PheLeuTyrPheValMetSer 377
1073 TTGCTCTATTCTTGTGTTCA 1093

RESULT 26
US-08-486-673B-62
; Sequence 62, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 1414
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(1240)

```

OTHER INFORMATION: C140 receptor, cDNA and deduced protein sequences
US-08-486-673B-62

Alignment Scores:
Pred. No.: 5,41e-48 Length: 1414
Score: 551.00 Matches: 118
Percent Similarity: 51.77% Conservative: 72
Best Local Similarity: 32.15% Mismatches: 139
Query Match: 25.80% Indels: 38
DB: 4 Gaps: 8

US-09-208-629F-3 (1-407) x US-08-486-673B-62 (1-1414)

Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
Db 74 CTGCTGGGGCCGCATCTCTAGAGAGCTTATTGGTAAGTTGATGTC-----181

Qy 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
Db 134 ACCAATAGATCTCTAAAGGAAGAGCTTATTGGTAAGTTGATGTC-----181

Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
Db 182 -----ACATCCACGTC-----ACTGGAAGAGGATTACAGTT 214

Qy 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
Db 215 GAA-----ACAGTCTTTCTGGGATGAGTTTCTGCATCT 250

Qy 102 TyrLeuArgSerSerLeuThrGlnValIleProAlaIleValIleLeuLeuPheVal 121
Db 251 GTCCTCGTGGAAACTGACCATGCTCTCTCCATTCATGCTACACAAATTGTTGCG 310

Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
Db 311 GTGGGTTTGCACGTAACGCGATGCGCTATGGCTCTTCTTTCCGAACCTAAGAGAAG 370

Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaLeuLeuPheCysValThr 160
Db 371 CACCTCTGTGATTATGAGGCAATCTGGCCCTTGGCTGACCTCTCTCTGTCATCGG 430

Qy 161 LeuProPheLysIleAlaLysHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 431 TTCCCTTGAAGATTGCTATCATACATGCAACCACTGGATTATGGGAAGCTCT 490

Qy 181 CysArgIleThrThrValPheThrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
Db 491 TGTAAATGTCATTATGGCTTTTCTATCGCAACATGACTGTTCCATCTCTTCATGACC 550

Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 551 TGCCTCAGTGTGCAGAGATTATGGTATCTGTGAACCCCATGGGGCACTCCAGG---AAG 607

Qy 221 LysArgSerSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 608 AAGCAAAACATTGCAATGGCATCTCCCTGCAATATGGCTGCTGCTGCTGCTGCTCACC 667

Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
Db 668 ATCCCTTTGTATGTCGTGAAGCAGACCATCTTCTATCTGCTGCTGCAACATCAGCACCTGT 727

Qy 261 HisAspValValAspAlaCysGluSerProSerPhe-----ArgPheTyr 276
Db 728 CATAGTGTGTTG-----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772

Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCys 296
Db 773 TACTTCCTCTCTGGCCATGGGGTCTTCTGTTCCAGCCTCTCTCAGACCTCTGCC 832

Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
Db 833 TATGTGCTGATGATCAGAAATGTCGATCTCTCTGCCATGATGAAACTCAGAGAAGAAA 892

Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAla 330
Db 893 AGGAAGAGGGCCATCAACTCATGTCACCTGCTGGGCACTGATCTGATCTGCTCACT 952

Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
Db 953 CCTAGTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012

Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
Db 1013 GTCATGCGCTGATACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072

Qy 371 PheLeuTyrPheValMetSer 377
Db 1073 TTGTCATACTTTGTTTCA 1093

RESULT 27
US-08-313-553-12
Sequence 12, Application US/08313553
Patent No. 5641650
GENERAL INFORMATION:
APPLICANT: TURNER, George J.
APPLICANT: BETLACH, Mary C.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
TITLE OF INVENTION: IN HALOBACTERIA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Walter H. Dreger
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/313,553
APPLICATION NUMBER: US/08/313,553
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/038,662
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-57669/WHI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1764 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: repeat region
LOCATION: 436..462
OTHER INFORMATION: /note= "Sequence encoding
Patent No. 5641650
OTHER INFORMATION: polypeptidic acid."
FEATURE:
NAME/KEY: misc feature
LOCATION: 463..465
OTHER INFORMATION: /note= "Codon encoding the
OTHER INFORMATION: N-terminal amino acid of the human thrombin
OTHER INFORMATION: receptor protein."
FEATURE:
NAME/KEY: misc feature

```

LOCATION: 1630..1632
OTHER INFORMATION: /note= "Codon encoding the
OTHER INFORMATION: C-terminal amino acid of the human thrombin
OTHER INFORMATION: receptor protein."
FEATURE:
NAME/KEY: repeat region
LOCATION: 1633..1650
OTHER INFORMATION: /note= "Sequence encoding
Patent No. 5641650
OTHER INFORMATION: polyhistidine."
FEATURE:
NAME/KEY: misc feature
LOCATION: 648..656
OTHER INFORMATION: /note= "Deleted AlwNI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc feature
LOCATION: 893..898
OTHER INFORMATION: /note= "Deleted PstI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc feature
LOCATION: 1301..1309
OTHER INFORMATION: /note= "Deleted AlwNI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc feature
LOCATION: 1394..1402
OTHER INFORMATION: /note= "Deleted AlwNI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc signal
LOCATION: 374
OTHER INFORMATION: /note= "RNA start site."
FEATURE:
NAME/KEY: mutation
LOCATION: replace(1671, "")
OTHER INFORMATION: /note= "C to T mutation removes
OTHER INFORMATION: AlwNI site."
FEATURE:
NAME/KEY: CDS
LOCATION: 376..1650
FEATURE:
NAME/KEY: misc feature
LOCATION: 376..414
OTHER INFORMATION: /note= "Bacteriorhodopsin
OTHER INFORMATION: pre-sequence."
FEATURE:
NAME/KEY: terminator
LOCATION: 1672..1674
OTHER INFORMATION: /note= "Bacteriorhodopsin stop
OTHER INFORMATION: codon."
FEATURE:
NAME/KEY: terminator
LOCATION: 1651..1653
OTHER INFORMATION: /note= "Thrombin stop codon."
US-08-313-553-12

Alignment Scores:
Pred. No.: 9.55e-44 Length: 1764
Score: 511.50 Matches: 122
Percent Similarity: 52.07% Conservative: 79
Best Local Similarity: 31.61% Mismatches: 148
Query Match: 23.95% Indels: 37
DB: 1 Gaps: 14

US-09-208-629F-3 (1-407) x US-08-313-553-12 (1-1764)

Qy 19 IleLeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThr-----ValCys 35
Db 355 ATCTCGTTAGTACTGTTGGCATGTTTGGCAACAGCAGGTGGAGGGGATCG 414
Qy 36 GlnSerGlyIleAsnVal-----SerAspAsnSerAlaLysProThrLeuThr 51

```

	:::	:::		
Db	415	CAGGCCAGATCCAGCGCTGGACTACAAAGGACGATGATGACGTTCGACGCCACCTTTAGAT	474	
Qy	52	IleYserPhe--AenGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAsp	70	
Db	475	CCCGGTCAITTTCTCAGGAACCCCAATGATAATATGAACCATTTTGGGAGGATGAG	534	
Qy	71	IleGluGlyTrpThrGlyAlaThr-----ThrThrIleyysAlaGluCysePro	86	
Db	535	GAGAAAATAAGAAATGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGACGAGTCT	594	
Qy	87	GluaspSerileSerThrLeuHisValAsnAsnAlaThrIleglyTyrrLeuArgSerSer	106	
Db	595	CTTCAAACAACACTTCTGCATTCACTCTCAGAAGATGCCCTCCGGATATTTGACCAGCGTCC	654	
Qy	107	LeuSerThrGlnValIleProAlaIleTyrrIleLeuLeuPheValValGlyValProSer	126	
Db	655	TGGCTGNACACTTTGTGCCCATCTGTGTACACCGGAGTGTGTGTAGTCACGCTCCCAC	714	
Qy	127	AsnIle-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal	144	
Db	715	AACATCATGCCCATCGTTGTGTTCATCTCTGAAATGAAGGTCAAGAAG---CGGCGGTG	771	
Qy	145	IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPhyls	164	
Db	772	GTGTACATGCTGCACCTGGCCAGCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTAA	831	
Qy	165	IleAlaTyrrHisLeuAsnGlyAsnAsnTrpValPheGlyLuValMetCysArgIleThr	184	
Db	832	ATCAGCTATTACTTTTCCGGCAGTGATGGCAGTTTGGGTCTGAATGTGTGGTTCGTC	891	
Qy	185	ThrValValPheTyrrGlyAsnMetTyrrCysAlaIleLeuIleLeuThrCysMetGlyIle	204	
Db	892	ACTGCAGCATTTTACTCTAACATGTAGCCCTCTATCTTGTCTCATGACAGTCATAGG	951	
Qy	205	AsnArgTyrrLeuAlaThrAlaHisPro-----PheThrTyrrGlnLysLeuProLys	221	
Db	952	GACCGGTTTCTGGCTGTGTGTATCCCATCAGTCCTCTCTGGCGTACTCTGGGAAGG	1011	
Qy	222	ArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrrMetLeu	241	
Db	1012	GTITCTCTTCACT-----TGTGTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGT	1062	
Qy	242	PropheValIleLeuLysGlnGluTyrrHisLeuValHisSerGluIleThrThrCysHis	261	
Db	1063	CCTCTCGTCTCAAGAGCAAAACCATCCAGGTGCCCGGGCTCAACATCACTACTGTG	1122	
Qy	262	AspValValAspAlaCysGluSerProSerSerPheArgPheTyrrTyrrPheValSerLeu	281	
Db	1123	GATGTGCTCAATGAAACCCCTGTCTCGAAGGTACTATGCTCTACTACTTCTCAGCCT	1182	
Qy	282	AlaphePheGlyLeuIleProPheValIleIleIlePheCysTyrrThrThrLeuIle	301	
Db	1183	GCYTGTCTC---TTTTTGTGGCGCTGATCATTTCCACGGCTGTGTATGTCTATCAT	1239	
Qy	302	HisLysLeuLysSer-----LysAspArgIleTrpLeuGly	313	
Db	1240	CGATGCTTAGCTCTCCGAGTTGCCAACCGCAGCAAGAGTCCCGGGCTTTG-----	1293	
Qy	314	TyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsn	333	
Db	1294	TTCTGTGCAGCTGCT-----GTTTTCTGCATCTTCATCATTTGCTTCGGACCCACA	1347	
Qy	334	IleIleLeuValIleHisHisAlaAsnTyrrTyrrHis-----AsnThrAspSerLeu	351	
Db	1348	GTCTCTGTATGGCATTAAC---TCATCTCTTCTCACATCTTCACACACGAGGCTGCC	1404	
Qy	352	TyrPheMetTyrrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProphe	371	
Db	1405	TACTTTGCCTTACCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1464	
Qy	372	LeuTyrrPheValMetSer	377	

Db 1465 ATTACTATTACGCTCC 1482

RESULT 28

US-08-767-993-12

Sequence 12, Application US/08767993

Patent No. 6010885

GENERAL INFORMATION:

APPLICANT: TURNER, George J.

APPLICANT: BETLACH, Mary C.

TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES

TITLE OF INVENTION: IN HALOBACTERIA

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Walter H. Dreger

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/767,993

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Walter H.

REGISTRATION NUMBER: 24,190

REFERENCE/DOCKET NUMBER: A-57669/WH

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1764 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: repeat region

LOCATION: 435..462

OTHER INFORMATION: /note= "Sequence encoding

Patent No. 6010885

OTHER INFORMATION: polyaspartic acid."

FEATURE:

NAME/KEY: misc feature

LOCATION: 463..465

OTHER INFORMATION: /note= "Codon encoding the

OTHER INFORMATION: N-terminal amino acid of the human thrombin

OTHER INFORMATION: receptor protein."

FEATURE:

NAME/KEY: misc feature

LOCATION: 1630..1632

OTHER INFORMATION: /note= "Codon encoding the

OTHER INFORMATION: C-terminal amino acid of the human thrombin

OTHER INFORMATION: receptor protein."

FEATURE:

NAME/KEY: repeat region

LOCATION: 1633..1650

OTHER INFORMATION: /note= "Sequence encoding

Patent No. 6010885

OTHER INFORMATION: polyhistidine."

FEATURE:

NAME/KEY: misc feature

LOCATION: 648..656

OTHER INFORMATION: /note= "Deleted AlwNI restriction

OTHER INFORMATION: site."

FEATURE:

NAME/KEY: misc feature

LOCATION: 126

OTHER INFORMATION: /note= "Thrombin stop codon."

US-08-767-993-12

Alignment Scores:

Pred. No.: 9.55e-44 Length: 1764

Score: 511.50 Matches: 122

Percent Similarity: 52.07% Conservative: 79

Best Local Similarity: 31.61% Mismatches: 148

Query Match: 23.95% Indels: 37

DB: 3 Gaps: 14

US-09-208-629f-3 (1-407) x US-08-767-993-12 (1-1764)

QY 19 IleLeuIleuValAlaAlaGlyLeuLeuPheLeuProValThr-----ValCys 35

Db 355 ATCTCGTTAGGTACTGTTGTCATGTTGAGTTATTCACACAGCAGTGGAGGGGTATCG 414

QY 36 GlnSerGlyIleAsnVal-----SerAspAsnSerAlaLysProThrLeuThr 51

Db 415 CAGGCCAGATCCAGGCGCTGACACTACAGGACGATGATGACGTCGACGCCACCTTAGAT 474

QY 52 IleLysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAsp 70

Db 475 CCCCCTCATTTCTTCTCAGGAACCCCAATGATAATATGAACCAATTTTGGGAGGATGAG 534

QY 71 IleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysPro 86

Db 535 GAGAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGCTCT 594

QY 87 GluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSer 106

Db 595 CTTCAAAAACAACTTCTTCATTCATCAGAGATGCTCCGATATTGACCACTCC 654

QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValProSer 126

Db 655 TGGCTGACACTTTTGTGCCATCTGTGTACACGGAGTGTGTGTAGTCAGCCTCCCACTA 714
Qy 127 AsnIle-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 715 AACATCATGCCATCGTTGTGTTTCATCTCGAAATGAAGTCAAGAAG---CCGGCGGTG 771
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 772 GTGTACATGCTGCACCTCGGCAGCGCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTAA 831
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 832 ATACGCTATTACTTTCCCGCAGTATGCGAGTTTGGGTCTGAATTTGTCTGCTTCGTC 891
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
Db 892 ACTGCAGCATTTTACTGTAAATGATGACGCTCTATCTTGTCTCATGACAGTCATAAGCAT 951
Qy 205 AsnArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLys 221
Db 952 GACCGGTTTCTGGCTGTGTGTGTATCCCATGCTGCTCTCTCTCTCTCTCTCTCTCTCT 1011
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeu 241
Db 1012 GCTTCTCTCTCT-----TGCTGCGCATCTGGGCTTTGGCCATCGCAGGGGTAGT 1062
Qy 242 ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHis 261
Db 1063 CCTCTGCTCTCAAGAGCAAAACCATCCAGGTGCGCGGCTCAACATCACTACTGTCTAT 1122
Qy 262 AspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu 281
Db 1123 GATGTGCTCAATGAAACCTGCTCGAAGGCTACTATGCTCTACTTCTCAGCCTTCTCT 1182
Qy 282 AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIle 301
Db 1183 GCTGCTCTCT-----TTTTTGTGCGCTGATCATTTCCACGGTCTGTATGTCTATCAT 1239
Qy 302 HisLysLeuLysSer-----LysAspArgIleThrLeuGly 313
Db 1240 CGATGCTTAGCTTCTCGCAGTTGCCACCGCAGCAAGAAGTCCCGGCTTTG----- 1293
Qy 314 TyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsn 333
Db 1294 TTCTGTGCTGCT-----GTTTCTGCTATCTTCATCTTCTGCTGCTGCTGCTGCTGCT 1347
Qy 334 IleIleLeuValIleHisHisAlaAsnTyrTyrHis-----AsnThrAspSerLeu 351
Db 1348 GTCTCTGATGCGCAATTAC---TCATTCCTTTCTCACACTTCCACACAGAGGTGCC 1404
Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProphe 371
Db 1405 TACTTTGCTACCTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1464
Qy 372 LeuTyrPheValMetSer 377
Db 1465 ATTACTATTACGCTTCC 1482

RESULT 29
US-07-657-769B-68
; Sequence 68, Application US/07657769B
; Patent No. 5256766
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; TITLE OF INVENTION: RECOMBINANT RECEPTOR AND RELATED
; TITLE OF INVENTION: PHARMACEUTICALS
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IRELL & MANELLA
; STREET: 545 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/657,769B
FILING DATE: 19910219
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2000-0502.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 3480 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 225..1503
US-07-657-769B-68

Alignment Scores:
Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 1 Gaps: 14

US-09-208-629F-3 (1-407) x US-07-657-769B-68 (1-3480)

Qy 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTCTGTGTGCGCCCTGCTTC-----AGTCTGTGCGGCCCGCTGTG 284
Qy 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TCTGCCCGCACCGGCCCGCCAGCAATCAAAAGCAACAATGCCACTTAGATCCC 344
Qy 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCACTTTCTCTCAGGAACCCCAATGATAATATGAACCAATTTGGGAGGATGAGAG 404
Qy 72 GluGlyTrpThrGlyAlaThr-----ThrIleLysAlaGluCysProGlu 87
Db 405 AAAAATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGACAGTCTCTT 464
Qy 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAACAACCTCTGCATTCTCAGGAAGTCCCGGATATTTGACCACTCTCTGG 524
Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTGACACTCTTTGTCCCTCTGTGTACACGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 584
Qy 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGCCATCTGTGTGTCTTCATCTCGAAATGAAGTCAAGG---CCGGCGGTGGT 641
Qy 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTGACCTGCGCAGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 701
Qy 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 185

QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCACT-----TGTGGCCATCTGGCTTTGGCCATCGCAGGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 933 CTGCTCTCAAGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACTGTCTGATGAT 992
QY 263 ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCCCTGCTCGAAGGCTACTATGCTTACTTCTCAGCCTTCTCTGCT 1052
QY 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
Db 1053 GTCTTC---TTTTTGTGGCGCTGATCATTTCCACGGCTGTGTATGTGTATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTyrPheLeuGlyTyr 314
Db 1110 TGTCTTAGCTTCTCCGAGTGTGCAACCGCAGCAAGAGTCCCGGGCTTTG-----TTC 1163
QY 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 1164 CTGTCAGCTCT-----GTTTCTGCATCTTCATCATTTGCTTCGGACCCCAAGGTC 1217
QY 335 IleLeuValIleHisHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
Db 1218 CTCTGATTGCGATTAC---TCATTCTTCTCACACTTCCACACAGAGGCTGCCTAC 1274
QY 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeu 372
Db 1275 TTTGCTTACTCTCTCTGTGTGTGTGTCAGCAGCATAGCTCGTGCTACGCCCTTAATT 1334
QY 373 TyrPheValMetSer 377
Db 1335 TACTATTACGCTTCC 1349

RESULT 31

US-08-475-263-219
; Sequence 219, Application US/08475263
; Patent No. 5759994
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/475,263
; APPLICATION NUMBER: US/08/475,263
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELE: 90-4030
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-475-263-219
Alignment Scores:
Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 1 Gaps: 14
US-09-208-629f-3 (1-407) x US-08-475-263-219 (1-3480)
QY 20 LeuIleLeuValAlaAlaGlyLeuLeuPheProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTGCTGCTGGCCGCTCTTC-----AGTCTGTGCGCCGCTGTTG 284
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TCTGCCCGCACCCGGCCGCGCAGCCAGAAATCAAAAGCAACAATGCCACCTTAGATCCC 344
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCACTTTCTCTCAGGAACCCCAATGATAATATGAACCATTTTGGGAGGATGAGGAG 404
QY 72 GluGlyTyrThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
Db 405 AAAAATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGCTCTCTT 464
QY 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAAACAATCTCTCGATTCATCTCAGAAAGATGCTCCGGATATTGTGACCACTCTCG 524
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTCACACTCTTTGTCCTCATCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCAATAAC 584
QY 128 Ile-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGGCCATCGTTGTGTTCATCTCTGAAATGAAGGTCAAGAAG---CGGCGGTGGTG 641
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTCACCTGCGCCACGCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 701
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 702 AGCTATTACTTTTCGGCAGTGAATGGCAGTTTGGGTCTGAATTTGTGCTGCTGCTCACT 761
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsn 205
Db 762 GCAGCATTTTACTGTAAATGATGACGCTCTATCTCTCATGACAGTCATAAGCATGAC 821
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
Db 822 CGGTTTCTGGCTGTGTGTATCCCATGCGATGCTCTCTCTGCGGTACTCTCGGGAAGGCT 881
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCACT-----TGTCTGGCCATCTGGGCTTTGGCCATCCGAGGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 933 CTGCTCTCAAGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACTCTGTCATGAT 992
QY 263 ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCCCTGCTCGAAGGCTACTATGCTTACTTCTCTCAGCCTTCTCTGCT 1052

QY 283 PhePheGlyPheLeuLeuProPheValIleIlePheCysTyrThrThrLeuIleHis 302
Db 1053 GTCTTC---TITTTTGGCGCGTATTCATTCACGGTCTGTATGTGCTATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 1110 TGTCTTAGCTCTTCCGCGAGTTGCCAACCAGCAAGAGTCCCGGGCTTTG-----TTC 1163
QY 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 1164 CTGTCAGCTGCT-----GTTTTCTGCATCTTCATCATTTGCTTCGACCCCAACAGTC 1217
QY 335 IleLeuValIleIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
Db 1218 CTCCTGATGGCATTAC---TCATTCCTTCTCACACITTCACACAGAGGTGCTCTAC 1274
QY 353 PheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
Db 1275 TTTGCTTACCTCCTCTGTGTGTGTGTCAGCAGCATAGCTCGTGCATCGACCCCTAATT 1334
QY 373 TyrPheValMetSer 377
Db 1335 TACTATTACGCTTCC 1349

RESULT 32

US-08-485-886-219
; Sequence 219, Application US/08485886
; Patent No. 5798248
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
US-08-485-886-219

Alignment Scores:

Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 1 Gaps: 14
US-09-208-629F-3 (1-407) x US-08-485-886-219 (1-3480)
QY 20 LeuIleLeuValAlaAlaGlyLeuLeuProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTGCTGGTGGCGGCTGCTTC-----AGTCTGTGCGCGCGCTGTGTTG 284
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TTTGCCCGCACCGCGCGCGCGAGTCAAAAGCAACAAATGCCACCTTAGATCCC 344
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCACTTCTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAG 404
QY 72 GluGlyTrpThrGlyAlaThr-----ThrIleLysAlaGluCysProGlu 87
Db 405 AAAAATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAAGCATCTCTT 464
QY 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAACAACCTTCCTGCATTCATCTCAGAGATGCTCCGGATATTTTGACCACTCCTGG 524
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTCACACTCTTGTCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCAAC 584
QY 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGGCCATGTTGTGTTCATCTCTGAAATGAAGTCAAGAAG---CCGGCGGTGGTG 641
QY 146 PheHisThrAsnLeuAlaIleAlaLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTGCACCTGGCCACGGAGATGTGTGTGTCTGTCTCTCCCTTTAGATC 701
QY 166 AlaTyrHisLeuAsnGlyAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 702 AGCTATTACTTTTCCGGCAGTGTGGTCTGAATGTGTGCTCGTCACT 761
QY 186 valValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn 205
Db 762 GCAGCATTTTACTGTAAACATGTACGCTCTATCTCTCATGACAGTCAATGACATGAC 821
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
Db 822 CGGTTTCTGGTGTGTGTATCCATGTCAGTCCCTCTCTGGGTACTCTGGGAAGGCT 881
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCAC-----TGTCTGGCCATCTGGCTTTGGCCCATCCAGGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAsp 282
Db 933 CTGCTCTCAAGGAGCAACCATCCAGGTGCCGGCTCAACATCACTACTACCTCATGAT 992
QY 263 ValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCCCTGTGGAAGGCTACTATGCTCTACTACTTCTCAGCCTTCTCTGCT 1052
QY 283 PhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeuIleHis 302
Db 1053 GTCTTC---TTTTTGTGCGGTGATCAITTCACGGTCTGTATGTGTCTATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 1110 TGTCTTAGCTCTTCCGCGAGTTGCCAACCAGCAAGAGTCCCGGGCTTTG-----TTC 1163

RESULT 35

US-08-473-489A-219
; Sequence 219, Application US/08473489A
; Patent No. 6024936
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,489A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,184
; FILING DATE: 1991-11-07
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-473-489A-219

Alignment Scores:
Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 3 Gaps: 14

US-09-208-629f-3 (1-407) x US-08-473-489A-219 (1-3480)

QY 20 LeuIleuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTGCTGGTGGCCGCTCTTC-----AGTCTGTCGGCCGCTGTG 284
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TCTGCCCGCCCGCCGCGCCGAGCAATCAAGCAACAAATGACCTTAGATCCC 344
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCACTTCTCTCAGGAACCCCAATGATAAATATGACCACTTTTGGGAGGATGAGGAG 404
QY 72 GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
Db 72 GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87

Db 405 AAAAATGAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCACTCTCTT 464
QY 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAACAACTTCTCCTCATTCATCTCAGAAGATGCTCCGGATATTTTGACCACTCTCTG 524
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTGACACTCTTTGTGCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCACTAAC 584
QY 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGGCCATCTGTGTTCATCTCGAAATCAAGGTCAAGAG---CGGCGGTGGT 641
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTGCACCTGGCCACGAGATGTCTGTGTCTGTCTGCTCCCTTTAAGATC 701
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 702 AGCTATTACTTTTCCGGCAGTGTGTCAGTGTGGGTCTGAATTTGTGTCTGCTCACT 761
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn 205
Db 762 GCAGCATTTTACGTAAACATGACGCTCTATCTGTCTATGACAGTCATCAAGCATGAC 821
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
Db 822 CGGTTTCTGGTGTGTGTATCCCATGACAGTCCCTCTCTGGGTACTCTGGGAAGGCT 881
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCACT-----TGCTGGCCATCTGGCTTGGCCATCGCAGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerCluIleThrCysHisAsp 262
Db 933 CTCGCTCTCAAGGAGCAACCATCCAGTCCCGGGCTCAACATCACTACCTGTCTCATGAT 992
QY 263 ValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCTGTCTGNAAGGTACTATGCTACTACTTCTCAGCCTTCTCTGCT 1052
QY 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
Db 1053 GTCTTC---TTTTTGTGGCTGATCAATTCACGGTCTGTATGTGTCTATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 1110 TGTCTTAGCTCTTCCAGAGTTCACACCGCAGCAAGAGTCCCGGGCTTTG-----TTC 1163
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 1164 CTGTGAGTGTCT-----GTTTCTGCATCTTTCATCATTTGTCGACCCCAACAGCTC 1217
QY 335 IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
Db 1218 CTCCTGATTGGCATTAC---TCATTCTTCTCACACTTCCACACAGAGGCTGCCTAC 1274
QY 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeu 372
Db 1275 TTTCCTACTCTCTGTGTGTGTGTCAGAGCATTAAGCTCGTGATCGACCCCTTAATT 1334
QY 373 TyrPheValMetSer 377
Db 1335 TACTATTACGCTTCC 1349

RESULT 36

US-08-485-695-219
; Sequence 219, Application US/08485695
; Patent No. 6124101
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.


```

; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4895 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 176..1330
; OTHER INFORMATION:
;
US-09-479-130-1
Alignment Scores:
Pred. No.: 1,09e-42 Length: 4895
Score: 507.50 Matches: 111
Percent Similarity: 52.97% Conservative: 76
Best Local Similarity: 31.44% Mismatches: 151
Query Match: 23.76% Indels: 15
DB: 4 Gaps: 6

US-09-208-629f-3 (1-407) x US-09-479-130-1 (1-4895)
Qy 27 LeuLeuPheLeuProValThrValCysGlnSerGlyLeuValSerAspAsnSerAla 46
Db 188 CTGCTCTGTCGGCCCTGGTGCTG-----GGGTTCAGCCTGTCTGCGCGCACCCAG 238
Qy 47 LysProThrLeuThrLeuLys-----SerPheAsnGlyGlyProGlnAsnThrPheGlu 64
Db 239 ACCCCAGCCTTACGACGAGAGCGGAGCGGAGGTGGTGATGACAGCAGCGCCCTCA 298
Qy 65 GluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrLeuLysAlaGlu 84
Db 299 ATCTGCTGCTCCCCC-----CGCGGCTACCCAGGCCAAGTC----- 334
Qy 85 CysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArg 104
Db 335 TGTGCCAATGACAGT---GACACCTTGAGCTCCCGACAGCTCACGGGCACTGCTTCG 391
Qy 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyVal 124
Db 392 GCGTGGGTGCCACACAGCGGTGGTCCCGCCTCTATGGGTGGTGGTGGTGGGCTG 451
Qy 125 ProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 452 CCGGCCAATGGCTGGCGCTGTGGTCTGGCCACGCGGCACTCGGCTGGCTCCACC 511
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 512 ATGCTGCTGATGAACCTCGGACCTGCTGCTGCTGGCCCTGGCGCTGCCCGCGG 571
Qy 165 IleAlaTyrHisIleuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 572 ATCCCTTACACCTCGCTGCGCCAGCGCTGGCCCTTCGGGGAGGCGGCTGCCGCTGGCC 631
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
Db 632 ACGGCCCACTTATGTGTACATGATGGCTCACTGCTGCTGGCGCGCGCTGACGCGCTG 691
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 692 GATCGCTACCTGGCCCTGGTGACCCGCTCGGGGCGCGCCCTGCTGGCGCGCGCCTG 751
Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 752 GCCCTTGGACTCTGATGGCTGTGCTGCTCATGGCGCGCGCCCTGACCTGCGCCCTGACA 811
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 812 CTGACGGGAGACCTTCCGGCTGGCGCGCTCCGATCGCTGCTGCTGCCATGACGCGCTG 871

US-08-383-750-3
; Sequence 3, Application US/08383750
; Patent No. 5744301
; GENERAL INFORMATION:
; APPLICANT: Birkenbach, Mark
; APPLICANT: Kieff, Elliot
; TITLE OF INVENTION: Epstein Barr Virus Induced Genes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.,
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,750
; FILING DATE: Herewith
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fox, Samuel, L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0627.3300001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..1116
;
US-08-383-750-3
Alignment Scores:
Pred. No.: 1.79e-35 Length: 1643
Score: 431.50 Matches: 95
Percent Similarity: 52.82% Conservative: 64
Best Local Similarity: 31.56% Mismatches: 107
```

Query Match:	20.20%	Indels:	35
DB:	1	Gaps:	9
US-09-208-629F-3 (1-407) x US-08-383-750-3 (1-1643)			
Qy	102	TyrLeuArgSerLeuSerLeuThrGlnValIleProAlaIleTyrIleLeuLeuPheVal	121
Db	103	TATGCACATCAGCAGCAGCGCAGGATAGTATGCCTCGCATACAGCCTGGTCTCATC	162
Qy	122	ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIle	141
Db	163	ATTGGGCTGGTGGGAAACTTACTAGCCCTGGTGCATGTGTTCATGTTCAAACAGGAAAAAATC	222
Qy	142	--SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr	160
Db	223	AACCTTACCACCCCTTATTCAACAAATTTGGTGATTTCTGTATATACTTTTTTACACGGCT	282
Qy	161	LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet	180
Db	283	TTGCCTACACGAATAGCCTACTATGCAATGGCTTTGACTGGAGAATCGGATGCGCTTG	342
Qy	181	CysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr	200
Db	343	TGTAGGATAACTGGCGTAGTGTTTTACATCAACACATATGCGAGTGTGAACCTTTATGACC	402
Qy	201	CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro	220
Db	403	TGCCTGAGTATTGACCGCTTCATGCTGTGTGGCACCTCTACGCTACACAAGATAAAA	462
Qy	221	LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet	240
Db	463	AGGATTGAACATGCAAAAAGCGGTGTGCATATTGTCTCGAATTCATAGTATTTGCTCAGACA	522
Qy	241	LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle	257
Db	523	CTCCCACTCTCATCAACCCCTATGTCAAGCAGGAG-----GCTGAAAGG	567
Qy	258	ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr	277
Db	568	ATTACATGCATGGAGTATCCAAACTTTGAAGAAACTAAATCT-----CTTCCCTGGATT	621
Qy	278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr	297
Db	622	CTGCTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATATCATCTCATCTGCTAT	681
Qy	298	ThrThrLeuIleHisLys-----LeuLysSerLysAsp	308
Db	682	TCTCAGATCTGTCGAAACTCTTCAAGACTGCCAAACAAACCCACTCACTGAGAAATCT	741
Qy	309	ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCys	328
Db	742	GGTGTAAACAAAAGGCTCTCAACACAATATTCTTATTTATTTGTTGTTGTTCTCTGT	801
Qy	329	PheAlaProThrAsnIleIleLeuValIleHis-----HisAlaAsn	342
Db	802	TTACACCTTACCATGTTGCAATATTCAACATATGATTAAAGAGCTTCGTTTCTCTAAT	861
Qy	343	Tyr-----TyrTyrHisAsnThrAsp-----SerLeuTyrPheMetTyrLeuIle	357
Db	862	TTCTCTGGAATGTAGCCAAAGACATTCGTTCCAGATTTCTCTCACCTTT-----	909
Qy	358	AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer	377
Db	910	ACAGTATGCTGATGAACCTTCAATTGCTGCATGGACCCCTTTTATCTACTTCTTCGATGT	969
Qy	378	Lys	378
Db	970	AAA	972

Search completed: June 29, 2003, 10:16:58
Job time : 102.944 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model.

Run on: June 29, 2003, 07:26:54 ; Search time 192.264 Seconds
(without alignments)
4767.217 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136

Sequence: 1 TLTYXHPVAGSQDKMKIL.....AMARPLXRRPRDIWEDIHAW 407

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09208629/runat_24062003_101625_12912/app_query.fasta_1.1166
-DB-N Geneseq 101002 -QWMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09208629 @CGN 1 1 275 @runat_24062003_101625_12912 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002: *
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2121	99.3	1224	19 AAV07372	Mouse protease-act
2	2121	99.3	2409	21 AA250774	Human protease act
3	1443.5	67.6	1224	19 AAV07374	Human protease-act
4	1443.5	67.6	1830	21 AAF21430	Human low adenosin
5	1443.5	67.6	1830	21 AAA35308	Human adenosine re
6	1443.5	67.6	6546	21 AAF21433	Human low adenosin
7	1443.5	67.6	6546	21 AAA35311	Human adenosine re
8	1369	64.1	1102	19 AAV07375	Human protease-act
9	1300.5	60.9	1124	19 AAV07373	Mouse protease-act
10	584	27.3	2732	16 AAQ84559	Murine C140 recept
11	584	27.3	2732	17 AAQ84557	Murine C140 recept
12	574.5	26.9	1477	16 AAQ84557	Murine C140 recept
13	572.5	26.8	1477	17 AAQ84558	Murine C140 recept
14	563	26.4	1255	16 AAQ84558	Human C140 recepto
15	563	26.4	1255	17 AAQ84558	Human C140 recepto
16	563	26.4	1289	21 AA250773	Human protease act
17	563	26.4	1451	21 AAC60319	Human PAR-2 DNA.
18	551	25.8	1414	16 AAQ84560	Human C140 recept
19	551	25.8	1414	17 AAQ84560	Human C140 recept
20	514.5	24.1	1312	18 AAT93367	Mouse thrombin rec
21	511.5	23.9	1764	15 AAQ73590	Fragment of the hu
22	509	23.8	3472	20 AA232191	Human thrombin rec
23	509	23.8	3480	21 AA250771	Human thrombin rec
24	507.5	23.8	1534	20 AAQ90981	Human protease-act
25	507.5	23.8	1534	21 AAF21431	Human low adenosin
26	507.5	23.8	1534	21 AAA35309	Human adenosine re
27	507.5	23.8	4895	20 AA232747	Human protease-act
28	507.5	23.8	4895	21 AA250775	Human protease act
29	507.5	23.8	4925	22 AAH43632	Human PAR4 coding
30	502	23.5	2910	18 AAT62461	Thr-GPAL fusion ge
31	502	23.5	3182	21 AAF21432	Human low adenosin
32	502	23.5	3182	21 AAA35310	Human adenosine re
33	501	23.5	1278	24 ABK70889	Human cDNA encodin
34	500	23.4	2137	22 AAH51006	Human nGPCR5 codin
35	497	23.3	1955	21 AA299588	CDNA encoding a hu
36	494	23.1	1080	21 AAD01131	Human orphan G-pro
37	494	23.1	1080	21 AAA46032	Human G-protein co
38	494	23.1	1361	20 AAQ90979	Mouse protease-act
39	493	23.1	1080	22 AAF57649	Human G-protein co
40	493	23.1	1080	23 AAI64231	G-protein coupled
41	493	23.1	1192	21 AAA14828	CDNA encoding a hu
42	491	23.0	3480	13 AAQ28568	Human thrombin rec
43	485.5	22.7	1209	24 ABK70887	Human cDNA encodin
44	484.5	22.7	1116	24 ABK70888	Human cDNA encodin
45	471.5	22.1	2588	19 AAV40372	Mouse G-protein co

ALIGNMENTS

RESULT 1
AAV07372
ID AAV07372 standard; cDNA; 1224 BP.
XX
AC AAV07372;
XX
XX
DT 12-OCT-1998 (first entry)
DE Mouse protease-activated receptor 3 (PAR3) cDNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosin;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
XX
OS Mus sp.
XX

FH Key Location/Qualifiers
 FT CDS 51..1160
 FT /*tag= a
 XX WO9818456-A1.
 PN 07-MAY-1998.
 PD 29-OCT-1997; 97WO-US19732.
 PF 30-OCT-1996; 96US-0742440.
 PR (REGC) UNIV CALIFORNIA.
 PA Connolly A, Coughlin SR, Ishihara H;
 PI WPI; 1998-271905/24.
 XX P-PSDB; AAW51405.
 DR
 DR
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 38-39; 74pp; English.
 XX
 CC This cDNA clone codes for mouse protease-activated receptor 3
 CC (PAR3) (see AAW51405), a cell surface protein which is specifically
 CC activated by thrombin or a thrombin agonist, thereby activating
 CC signalling events such as phosphoinositide hydrolysis, calcium ion
 CC efflux and platelet aggregation. A cDNA library was established
 CC from rat platelets and amplified (primers given in AAV07376-77). A
 CC product was identified that was expected to encode a new G-protein
 CC coupled receptor, related to PAR1 or 2. This was used to isolate
 CC human and murine DNA by a combination of PCR and hybridisation
 CC techniques. Murine cDNA, murine genomic DNA (see AAV07373), human
 CC cDNA (see AAV07374) and human genomic DNA (see AAW51406) sequences are
 CC provided, as well as the murine and human (see AAW51406) PAR3
 CC proteins. Also claimed are vectors, host cells and an assay
 CC device. Host cells are used to screen compounds for their ability
 CC to act as agonists or antagonists of the effects of thrombin-PAR3
 CC interaction. Agonists are used to treat wounds, thrombosis,
 CC atherosclerosis, restenosis, inflammation and other thrombin
 CC activated disorders. Antagonists (see AAW51415-21) are used to
 CC control blood coagulation and thereby to treat heart attack and
 CC stroke. They also mediate inflammatory and proliferative responses
 CC to injury as occur in wound healing, atherosclerosis, restenosis,
 CC pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 1224 BP; 299 A; 341 C; 239 G; 345 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,15e-188 Length: 1224
 Score: 2121.00 Matches: 404
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 19 Gaps: 0
 US-09-208-629F-3 (1-407) x AAV07372 (1-1224)
 Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
 Db 3 ACTTTGTATACTTAACAACATCCTGTAGCGGGTCTCAGGACATCAAGATGAAATCCTT 62
 Qy 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
 Db 63 ATCTTGGTGTGAGCTGGGCTCTCTTCTGCGAGTCATCTGTTCCCAAGTGGCATAAAT 122
 Qy 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
 Db 123 GTTTCAGACAACTCAGCAACCCACCTTAATTAAGAGTTTAAATGGGGGTCCTCCAA 182
 Qy 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80

Db 183 AATACCTTTGAAGATTCCCACTTTCTGACATAGAGGCTGGACAGAGCCACCAACT 242
 Qy 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100
 Db 243 ATAAAGCGGAGTGTCCCGAGGACAGTATTTCAACTCTCCAGCTGAATATGTACCAT 302
 Qy 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
 Db 303 GGATACCTGAGAAGTTCCTTAACTACCAAGTGATACCTGCCATCTATATCTCTGCTT 362
 Qy 121 ValValGlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSer 140
 Db 363 GTGGTTGGTGTACCATCCAACTCGTACCCCTGTGAAACTCTCTCTTAAGACCAATCC 422
 Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 Db 423 ATCAGTCTGGTATCTTTTACACCAACCTGGGCATCGCAGATCTCTTTTCTGTGCACA 482
 Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnMetTyrCysAlaIleLeuLeuThr 180
 Db 483 CTGCCATTTAAGATCGCTTACCATCTCAATGGCAACAACCTGGGTATTTGGCGAGGTATG 542
 Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
 Db 543 TGCCGGATCACACGGTCTGTTTCTACGGCAACATGTACTGCGCTATCTGTATCTCACT 602
 Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 Db 603 TGATGGGATCAACCGCTACCTGGCCAGCGCTACCCCTTTACATACCAAGAGTGGCC 662
 Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMet 240
 Db 663 AAACGGAGCTTCTCTTGTCTCATGTGTGCATAGTGTGGTCAATGGTTTCTTATACATG 722
 Qy 241 LeuProPheValIleLeuLysGlnLysLeuValHisSerGluIleThrThrCys 260
 Db 723 CTGCCCTTTGTCTCTCTGAAGCAGGAGTAGTACCACCTCGTCCACTCAGAGATCACCCCTGC 782
 Qy 261 HisAspValValAspIleCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
 Db 783 CACGATGTCTGCACGCGTGGAGTCCCATCATCTCTCCGATTTCTACTACTTCGTCTCC 842
 Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 843 TTAGCATTTCTTGGGTCTCTCATCCCGTTTGTGATCATCATCTTCTGTACAGACTCTC 902
 Qy 301 IleHisLysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeu 320
 Db 903 ATCCACAACTTAAATCAAAGGATCGATATGGCTGGGTATCATCAAGGCGCTCTCTCCTC 962
 Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
 Db 963 ATCCTTGTGATTTTCAAAATTTGCTTTGGCCCCCACCACCAATCATCTCGTAATCCACCAT 1022
 Qy 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCys 360
 Db 1023 GCCAACTACTACTACCAATACCGACAGCTGTACTTTATGTATCTTATGTCTCTGTGC 1082
 Qy 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
 Db 1083 CTGGGAGCGCTGAATAGCTGCCTAGATCCATTCCTTTACTTTCTATCTGTCTGTAAGTGA 1142
 Qy 381 AspGlnLeuAsnPro***SerAlaMetAlaAatqProLeu***ArgProAatqArgAspIle 400
 Db 1143 GATCAGCTTAATCCTTTAGTCGGAATGGCAAGACCACCTTTAGAGACCAAGGAGATATC 1202
 Qy 401 TrpGluAspIleHisAlaTrp 407
 Db 1203 TGGGAAGACATACATCTCTGG 1223
 RESULT 2
 AAZ50774

ID AAZ50774 standard; DNA; 2409 BP.
 AC AAZ50774;
 XX
 DT 31-MAY-2000 (first entry)
 XX
 DE Human protease activated receptor-3 DNA.
 XX
 KW Human; PAR-3; protease activated receptor;
 KW antisense molecule; PAR antibody; cytostatic; therapeutic;
 KW metastatic tumour cell; placental implantation; invasive cell; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200008150-A1.
 XX
 PD 17-FEB-2000.
 XX
 PF 05-FEB-1999; 99WO-1100079.
 XX
 PR 07-AUG-1998; 98IL-0125698.
 XX
 PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 XX
 PI Bar-Shavit R;
 XX
 DR WPI; 2000-205706/18.
 XX
 PT Treating metastatic tumor cells useful for treating disorders involving
 PT placenta implantation in a female comprises administration of an
 PT antisense molecule complementary to an RNA sequence of a protease
 PT activated receptor protein -
 XX
 PS Example 3; Fig 10; 46pp; English.
 XX
 CC The patent discloses a method to treat metastatic tumour cells using
 CC an antisense molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisense molecules and
 CC antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a human PAR-3
 CC DNA used for producing antisense molecules for treating invasive
 CC cells.
 XX
 SQ Sequence 2409 BP; 598 A; 652 C; 454 G; 703 T; 2 other;
 Alignment Scores:
 Pred. No.: 7.63e-188 Length: 2409
 Score: 2121.00 Matches: 404
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 21 Gaps: 0

US-09-208-629F-3 (1-407) x AAZ50774 (1-2409)

QY 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
 |||||
 DB 105 ACTTTGATACTTAACAAACATCTCTAGCCGGTCTCAGGACATCAAGATAAATCCTT 164
 |||||

QY 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
 |||||
 DB 165 ATCTTGGTGCAGCTGGGCTGCTTCTGCCAGTCACTGTTGCCAAAGTGCATAAAT 224
 |||||

QY 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyProGln 60
 |||||
 DB 225 GTTTCAGCAACATCAGCAAGCCAAACCTTAACATTAAAGAGTTTAAATGGGGTCCCAA 284
 |||||

QY 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80
 |||||
 DB 285 AATACCTTTGAGAAATTCCTCCACTTTCTGACATAGAGGGCTGGACAGGACCAACCAACT 344
 |||||

QY 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIle 100
 |||||

DB 345 ATAAAAGCGAGTGTCCTCCGAGGACAGTATTTCACCTCTCCACGTGAATAATGTACCATA 404
 |||||

QY 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
 |||||

DB 405 GGATACCTGAGAAGTTCTTTAAGTACCCAAGTGATACCTGCCATCTATATCTCTGCTGTTT 464
 |||||

QY 121 ValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer 140
 |||||

DB 465 GTGGTTGGTGTTACCATCCAAACATCGTGACCTGTGGAAACTCTCTTAAGGACCAATCC 524
 |||||

QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 |||||

DB 525 ATCAGTCTGGTCACTCTTTCACCAACCTGGCCATCGCAGATCTCCTTTCTGTGTGACA 584
 |||||

QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 |||||

DB 585 CTGCCATTTAAGATCGCTTACCATCTCAATGGCAACAACCTGGGTATTTGGCGAGGTCATG 644
 |||||

QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
 |||||

DB 645 TGCCGGATCACCAAGCGTCTTTTCTACGGCAACATGCTACTGGCTATCTCTGATCTCTCACT 704
 |||||

QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 |||||

DB 705 TGCATGGCATCAACCGCTACCTGGCCAGCGCTCACCTTTCACATACCAAGAGCTGCC 764
 |||||

QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 |||||

DB 765 AAACGAGCTTCTCTTGTCTCATGTGTGCATAGTGTGGTCAITGGTTTCTTATACATG 824
 |||||

QY 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
 |||||

DB 825 CTGCCCTTTGTCACTCTGAGCAGGAGTACCACCTCTGCTCCACTCAGAGATCACCACTGC 884
 |||||

QY 261 HisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
 |||||

DB 885 CACGATGTCTGCACGCGTGCAGAGTCCCATCATCTTCCGATTTCTACTACTTCTGCTCC 944
 |||||

QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 |||||

DB 945 TTAGCATTTCTTGGGTCTCTCATCCCGTTTGTGATCATCATCTTCTGTACAGACTCTC 1004
 |||||

QY 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
 |||||

DB 1005 ATCCCAAACTTAAATCAAAGGATCGGATATGCTGGGTACATCAAGCCGCTCTCTCCTC 1064
 |||||

QY 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
 |||||

DB 1065 ATCCTTGTGATTTTCACAATTTGCTTTGGCCCCCACCACATCATCTCGTAATCCACCAT 1124
 |||||

QY 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleIleAlaLeuCys 360
 |||||

DB 1125 GCCAACTACTACTACCACAATACCGACAGCTGTACTTATGTATCTTATGTCTGTGTGC 1184
 |||||

QY 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
 |||||

DB 1185 CTGGGGAGCTGNAATAGCTGCCTAGATCCATTCCTTACTTGTCTGTCGAAGTTGTA 1244
 |||||

QY 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
 |||||

DB 1245 GATCAGCTTAATCTCTAGTCGGCAATGGCAAGACCACCTTTAGAGACCAAGGAGATATC 1304
 |||||

QY 401 TrpGluAspIleHisAlaTrp 407
 |||||

DB 1305 TGGGAAGACATACATGCTTGG 1325
 |||||

RESULT 3
 AAV07374
 ID AAV07374 standard; cDNA; 1224 BP.
 XX
 AC AAV07374;
 XX

DT 12-OCT-1998 (first entry)
 XX Human protease-activated receptor 3 (PAR3) cDNA.
 DE
 XX
 KW Protease-activated receptor 3; PAR3; thrombin receptor; human;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 58..1182
 FT /*tag= a
 XX
 PN W09818456-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 29-OCT-1997; 97WO-US19732.
 XX
 PR 30-OCT-1996; 96US-0742440.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 XX Connolly A, Coughlin SR, Ishihara H;
 XX
 DR WPI; 1998-271905/24.
 DR P-PSDB; AAW51406.
 XX
 XX DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 40-41; 74pp; English.
 XX
 CC This cDNA clone codes for mouse protease-activated receptor 3
 CC (PAR3) (see AAW51406), a cell surface protein which is specifically
 CC activated by thrombin or a thrombin agonist, thereby activating
 CC signalling events such as phosphoinositide hydrolysis, calcium ion
 CC efflux and platelet aggregation. A cDNA library was established
 CC from rat platelets and amplified (primers given in AAV07376-77). A
 CC product was identified that was expected to encode a new G-protein
 CC coupled receptor, related to PAR1 or 2. This was used to isolate
 CC human and murine DNA by a combination of PCR and hybridisation
 CC techniques. Murine cDNA (see AAV07372), murine genomic DNA (see
 CC AAV07373), human cDNA and human genomic DNA (see AAV07375) sequences
 CC are provided, as well as the murine (see AAW51405) and human PAR3
 CC proteins. Also claimed are vectors, host cells and an assay
 CC device. Host cells are used to screen compounds for their ability
 CC to act as agonists or antagonists of the effects of thrombin-PAR3
 CC interaction. Agonists are used to treat wounds, thrombosis,
 CC atherosclerosis, restenosis, inflammation and other thrombin
 CC activated disorders. Antagonists (see AAW51415-21) are used to
 CC control blood coagulation and thereby to treat heart attack and
 CC stroke. They also mediate inflammatory and proliferative responses
 CC to injury as occur in wound healing, atherosclerosis, restenosis,
 CC pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 1224 BP; 303 A; 316 C; 221 G; 384 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 4,51e-125 Length: 1224
 Score: 1443.50 Matches: 264
 Percent Similarity: 83.74% Conservative: 45
 Best Local Similarity: 71.54% Mismatches: 59
 Query Match: 67.58% Indels: 1
 DB: 19 Gaps: 1
 US-09-208-629F-3 (1-407) x AAV07374 (1-1224)
 Qy 11 GlySerGlnAspIleLysMetLysIleLeuValAlaGlyLeuLeuPheLeu 30

Db
 40 GGGACTCAGGTCAATCAAAATGAAAGCCCTCATCTTTGACAGCTGCTGGCCCTCCTGCTTCTG 99
 Qy
 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
 Db
 100 TTGCCCACTTTTGTGAGAGTGGCAATGATAAACAACATGGAACAACTGGCAAGCAACC 159
 Qy
 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
 Db
 160 TTACCCATTAGACCTTTCTGAGAGCTCCCAAAATCTTTTGAAGAGTTCCCTTTTCT 219
 Qy
 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
 Db
 220 GCCTTGGAGGCTGGACAGGACGACGATTAAGTAAAAAATTAAGTGCCTGCAAGAAAGT 279
 Qy
 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
 Db
 280 GCTTCACATCTCCATGTGAAATGCTACCATGGGTACCTGACCAAGCTCCCTTAAGTACT 339
 Qy
 110 GlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleVal 129
 Db
 340 AAACGTATACCTGCCATCTACCTCTGGTGTGTGTAGTTGGTGTCCCGGCCAATGCTGTG 399
 Qy
 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
 Db
 400 ACCCTGTGATGCTTTCTTCAGGACCAAGATCATCTGTACCACTGTATTCTACCAAC 459
 Qy
 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
 Db
 460 CTGGCCATTGCAGATTTCTTTTGTGTGTACATTTGCCCTTTAAGATAGCTTATCATCTC 519
 Qy
 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
 Db
 520 AATGGAAACAATGGGTATTTGGAGAGGTCCTGTGCCGGGCCACCAAGATCATCTCTAT 579
 Qy
 190 GlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
 Db
 580 GGCAACATGCTACTGCTCCATTCTGCTCTGCTGCATCAGCATCAACCGTACCTGGCC 639
 Qy
 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
 Db
 640 ATCGTCCATCTTTCACCTACCGGGCCCTGCCCAAGCACACCTATGCTTGGTAACATGT 699
 Qy
 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
 Db
 700 GGACTGCTGGCAACAGCTTTCTTATATATGCTGCCATTTTTCATCTGAAGCAGGAA 759
 Qy
 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
 Db
 760 TATTATCTTGTTCAGCCAGACATCACCATGCTGCATGTTTTCATCTGAAGCAGGAA 819
 Qy
 270 ProSerSerPheArgPheTyrThrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
 Db
 820 TCATCTCCCTTCCAACTATTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 879
 Qy
 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAspArg 309
 Db
 880 TTTGCTCTTATCATCTATGCTATGAGGACATCATCCGACACCTTATGATGATGATCAT 939
 Qy
 310 IleTrpLeuGlyTyrIleLysAlaValLeuLeuValIleLeuValIlePheThrIleCysPhe 329
 Db
 940 AGATGTTGTGTGTATGTTAAGCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 999
 Qy
 330 AlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
 Db
 1000 GCTCCAAGCAATATTATTCTTATTATTCACCACTGCTAACTACTACTACTACTACTACT 1059
 Qy
 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
 Db
 1060 GGCCTTATATTTATATATCTCATAGCTTTGCTGGTAGTCTCTTATAGTCTCTAGAT 1119
 Qy
 370 ProPheLeuTyrPheValMetSerLys 378

Db 1120 CCATTCCTTTATTTCTCATGTCAAAA 1146

RESULT 4

AAF21430

ID AAF21430 standard; DNA; 1830 BP.

XX AC AAF21430;

XX AC AAF21430;

XX 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2997.

XX

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

XX human; airway disorder; bronchoconstriction; lung inflammation;

XX surfactant depletion; respiratory; bronchodilator; antiinflammatory;

XX immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;

XX respiratory obstruction; pulmonary obstruction; impeded respiration;

XX surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;

XX respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;

XX pulmonary hypertension; emphysema; pulmonary transplantation rejection;

XX chronic obstructive pulmonary disease; pulmonary infection; bronchitis;

XX cancer; ss.

XX OS Homo sapiens.

XX WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not

XX trigger adenosine receptors during metabolism, useful e.g. for treating

XX cancers and respiratory obstructions -

XX Disclosure; Page 1414-1415; 1592pp; English.

XX

XX The present invention describes low adenosine (A) content antisense

XX oligonucleotides and compositions (I) comprising them. In the antisense

XX oligonucleotides the A is replaced by a 'Universal' or alternative base.

XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,

XX immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.

XX The antisense oligonucleotides and (I) can be used to down-regulate the

XX expression and or activity of target polypeptides associated with

XX lung/respiratory disorders and malignancies, such as stimulating and

XX activating peptide factors and transmitters, transcription factors,

XX immunoglobulins and antibodies, antibody receptors, cytokines and

XX chemokines, endogenously produced specific and non-specific enzymes,

XX binding proteins, adhesion molecules and their receptors, cytokine and

XX chemokine receptors, adenosine receptors, bradykinin receptors, central

XX nervous system (CNS) and peripheral nervous and non-nervous system

XX receptors, CNS and peripheral nervous and non-nervous system peptide

XX transmitters, defensins, growth factors, vasoactive peptides and

XX receptors, binding proteins and malignancy associated proteins. The

XX antisense oligonucleotides may be used in this way to treat disorders

XX including respiratory obstruction (especially pulmonary obstruction

XX and/or bronchoconstriction) and/or lung inflammation, allergies)

XX and/or surfactant hypoproduction which are associated with a disease or

XX condition selected from pulmonary vasoconstriction, inflammation,

XX allergies, asthma, impeded respiration, respiratory distress syndrome

XX (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary

XX hypertension, emphysema, chronic obstructive pulmonary disease (COPD),

XX pulmonary transplantation rejection, pulmonary infections, bronchitis,

XX and/or cancer. AAF18434 to AAF21543 represent human polynucleotide

CC fragments and antisense oligonucleotides used in the exemplification of

CC the present invention.

XX

SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;

Alignment Scores:

Pred. No.: 7.64e-125 Length: 1830

Score: 1443.50 Matches: 264

Percent Similarity: 83.74% Conservatives: 45

Best Local Similarity: 71.54% Mismatches: 59

Query Match: 67.58% Indels: 1

DB: 21 Gaps: 1

US-09-208-629f-3 (1-407) x AAF21430 (1-1830)

QY 11 GlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeuLeuPheLeu 30

DB 127 GGGACTCAGGTCAATCAAAATGAAAGCCCTCACTTTGGAGCTGCTGGCTCTCTCTG 186

QY 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49

DB 187 TTGCCCACTTTTGTGCAGAGTGCATGGAAATGATACAAACAACCTTGGCAAGCCAACC 246

QY 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69

DB 247 TTACCCATTAAGACCTTTCTGTCGAGCTCCGCCAAATCTTTTGAAGAGTTCCCTTTTCT 306

QY 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89

DB 307 GCCTTGGAAAGCTGGACAGGACCCACGATTAAGTAAATTAAGTCCCTTGAAGAAAGT 366

QY 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109

DB 367 GCTTCACATCTCCATGTGAAATGCTACCATGGGTACTGACCCAGCTCTTAAAGTACT 426

QY 110 GlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleVal 129

DB 427 AAACCTGATACCTGCCATCTACCTCTCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 486

QY 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149

DB 487 ACCCTGTGATGCTTTCTTTCAGGACCATCACTGTACCATCTGATTTCTACACCAAC 546

QY 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169

DB 547 CTGGCCATTGCAGATTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 606

QY 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189

DB 607 AATGGGAACAACCTGGGTATTTGGAGAGGTCTGTGTGGCGGGCCACACAGTCATCTCTAT 666

QY 190 GlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAla 209

DB 667 GGCAACATGTACTGTCTTCATTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTG 726

QY 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229

DB 727 ATCGTCCATCTTTACCTACCGGGGCTGCCCAACACACCTATGCTTGTGTAACATGT 786

QY 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249

DB 787 GGACTGTGTGGGCAACAGTTTCTTATATATATGTCTGCCATTTTTCATCTACTGAAGCAGGA 846

QY 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269

DB 847 TATTATCTTGTTCAGCAGACATCACCATCTGCATGTATGTCACACCTTCCGAGTCC 906

QY 270 ProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289

DB 907 TCATCTCCCTTCAACTCTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 966

QY 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAspArg 309

Db 967 TTTGTGCTTATCATCTACTGCTATGCGAGCCATCATCGGACACTTAATGATCATACGATCAT 1026
 Qy 310 IletirpLeuGlyTyrIleLysAlaValLeuLeuLeuLeuVallePheThrIleCysPhe 329
 Db 1027 AGATGGTTGGTATGTAGGGGAGTCTCTCTATCTCTGTTGATTTTACCATTTGCTTT 1086
 Qy 330 AlaProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
 Db 1087 GCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACAACACTGAT 1146
 Qy 350 SerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
 Db 1147 GGCTTATATTTTATATATCTCATAGCTTTGTGCTGGGTAGTCTTATATAGTCTTAGAT 1206
 Qy 370 ProPheLeuTyrPheValMetSerLys 378
 Db 1207 CCATTCCTTATTTTCTCAATGTCNAAA 1233
 RESULT 5
 ID AAA35308
 XX AAA35308 standard; DNA; 1830 BP.
 AC AAA35308;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:182.
 XX
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200009525-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17712.
 XX
 PR 03-AUG-1998; 98US-0095212.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-205971/18.
 XX
 PT New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers -
 XX
 PS Disclosure; Page 1329; 1343pp; English.
 CC
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasise to the lungs, including

CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA3213 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX
 SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;
 Alignment Scores:
 Pred. No.: 7,64e-125 Length: 1830
 Score: 1443.50 Matches: 264
 Percent Similarity: 83.74% Conservative: 45
 Best Local Similarity: 71.54% Mismatches: 59
 Query Match: 67.58% Indels: 1
 DB: 21 Gaps: 1
 US-09-208-629F-3 (1-407) x AAA35308 (1-1830)
 Qy 11 GlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeuPheLeu 30
 Db 127 GGGACTCAGTCAATCAAAATGAAAGCCCTCATCTTTGACAGCTGCTGCCTCTGCTTCTG 186
 Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlalysProThr 49
 Db 187 TTGCCCATCTTTTGTGTCAGAGTGCGCATGGAATAATGATACAAACAACTTGGCAAGCAACC 246
 Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
 Db 247 TTACCCATTAAAGACCTTTCTGTGAGCTCCCCAAATTCCTTTTGAAGAGTTCCCTTTTCT 306
 Qy 70 AspIleGluGlyTyrThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
 Db 307 GCCTTGGAGGCTGGACAGGAGCCAGCATTTACTGTATAAATAATTAAGTGCCTGAAGAAAGT 366
 Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
 Db 367 GCTTCACATCTCCATGTGAAAATGCTACCATGGGTACCTGACCAAGCTCTTTAAGTACT 426
 Qy 110 GlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleVal 129
 Db 427 AAACGTATACCTGCCATCTACCTCTCTGGTGTGTGGTGTCCCGCCCAATGCTGTG 486
 Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
 Db 487 ACCCTGTGGATGCTTTTCTTCAGGACCAGATCATCTGTACCATCTATTCTACCAAC 546
 Qy 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
 Db 547 CTGGCCATTGCGAGATTTCTTTTGTGTATACATTCCTCTTAAAGTAGTATTATCATCTC 606
 Qy 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
 Db 607 AATGGGAACAACATGGGTATTTGGAGAGGTCTGTGCGGGCCACCAAGTCATCTTCTAT 666
 Qy 190 GlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
 Db 667 GGCAACATGATGCTCTCCATCTGCTCTTGTGCTGATCATGAGCATCAACCGCTACCTGGCC 726
 Qy 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
 Db 727 ATCGTCCATCTCTTACCTACCGGGGCTGCCCAAGCACACCTATGCTTGGTAAACATGT 786
 Qy 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
 Db 787 GGACTGTGTGGCAACAGATTCTTATATATGCTGCCATTTTTCATCTGAAGACAGGAA 846

Qy 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
Db 847 TATTATCTGTTCAGCCAGACATCACACCTGCCATGCTTCAACACTTGGAGTCC 906
Qy 270 ProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
Db 907 TCATCTCCCTCCCACTCTATTACTTCATCTCCTTGGCATCTTTGGATTCTTAATTCGA 966
Qy 290 PheValIleIlePheCysTyrThrThrLeuIleHisLysLysSerLysAspArg 309
Db 967 TTTGTGCTTATCATCTACTGCTATGCAGCCATCATCCGGACACTTAATGCATCATCAT 1026
Qy 310 IleTrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPhe 329
Db 1027 AGATGGTGTGGTATGTTAAGGCGAGTCTCTCATCTCTGTGATTTTACCATTTGCTTT 1086
Qy 330 AlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
Db 1087 GCTCCAGCAATATATCTTATTTATTCACCATGCTAACTACTACTAACAACACTGAT 1146
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1147 GGCTTATATTTATATATCTCATAGCTTTGGCTGGGTAGTCTTATAGTTGCTTAGAT 1206
Qy 370 ProPheLeuTyrPheValMetSerLys 378
Db 1207 CCATTCTCTTATTTCTCATGTCAAAA 1233

RESULT 6

AAF21433
ID AAF21433 standard; DNA; 6546 BP.
XX AAF21433;
AC
XX
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #3000.
XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impaired respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
OS Homo sapiens.
XX
XX W0200062736-A2.
PN
XX
PD 26-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US08020.
XX
XX 06-APR-1999; 99US-0127958.
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX (NYCE/) NYCE J W.
XX
XX Nyce JW;
XX
XX WPI; 2000-679539/66.
XX
XX Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
XX Disclosure; Page 1416-1418; 1592pp; English.
PS
XX The present invention describes low adenosine (A) content antisense
CC

CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with and
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impaired respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 6546 BP; 1603 A; 1691 C; 1500 G; 1751 T; 1 other;

Alignment Scores:

Pred. No.: 4, 05e-124 Length: 6546
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 21 Gaps: 1
US-09-208-629F-3 (1-407) x AAF21433 (1-6546)

Qy 11 GlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeuLeuPheLeu 30
Db 127 GGGACTCAGTCTCAATAAAGGCCTCATCTTTGACGCTGCTGCTGCTGCTGCTGCTG 186
Qy 31 ProValThrValCysGlnSerGlyIle--AsnValSerAspAsnSerAlaLysProThr 49
Db 187 TTCCCCACTTTTGTGAGAGTGGCAATGATAACAACTTGGCAAGCCCAACC 246

Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 247 TTACCCATTAAAGACCTTCGTGGAGCTCCCAAAATCTTTTGAAGAGTTCCCTTTTCT 306

Qy 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 307 GCCTTGGAAAGGCTGGACAGGACGACGATCTACTGTAAAAAATTAAGTGCCTGAAGAA 366

Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 367 GCTTCACATCTCCATGTGAAATGTACCATGGGTACCTGACCATGCTTAAAGTACT 426

Qy 110 GlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleVal 129
Db 427 AAATGATACCTGCCATCTACCTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 486

Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 487 ACCCTGTGGATGCTTTTCTTCAGGACGACATCCATCTGTACCATCTGTATCTACACCA 546

Qy 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
Db 547 CTGCCATTGCAGATTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 606

QY 170 AsnGlyAsnAsnTrpValPheGluValMetCysArgIleThrThrValValPheTyr 189
Db 607 AATGGGAACAACTGGGTATTTGGAGAGTCTGTGCGGGCCACACAGTCATCTCTAT 666
QY 190 GlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
Db 667 GGCAACATGATCTGCTCAATCTGCTTGGCTGCATCAGCATCAACCGCTACCTGGCC 726
QY 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
Db 727 ATGCTCATCTTTCACTACCGGGCTGCCCCAAGCACACATGCTTGGTAACATGT 786
QY 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
Db 787 GGACTGTGTGGCAACAGTTTCTTATATATGCTGCCATTTTCATCTGAAGCAGGAA 846
QY 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
Db 847 TATTATCTTGTTCAGCCAGACATCACCCCTGCCATGATGTCACAACTTGGAGTCC 906
QY 270 ProSerSerPheArgPheTyrTyrPheValSerLeuAlaPheGlyPheLeuIlePro 289
Db 907 TCATCTCCCTCCCAACTTATTACTTCATCTCCTTGGCATTTCTTGATTTCTTAATCCA 966
QY 290 PheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAspArg 309
Db 967 TTGTGCTTATCATCTACTGATGACCATCATCCGACATTAATGATAGCATCAT 1026
QY 310 IleTrpLeuGlyTyrIleLeuAlaValLeuLeuLeuValIlePheThrIleCysPhe 329
Db 1027 AGATGTTGTGTATGTTAAGCGAGTCTCTCATCTCTGTGATTTTACCATTTGCTTT 1086
QY 330 AlaProThrAsnIleLeuValIleHisHisAlaSerTyrTyrHisAsnThrAsp 349
Db 1087 GCTCCAAAGCAATATATTTCTTATTATTCACCATGCTAACTACTACTACAAACACTGAT 1146
QY 350 SerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1147 GGCTTATTTTATATATCTCATAGCTTTGTGCTGGGTAGTCTTATATAGTTGCTAGAT 1206
QY 370 ProPheLeuTyrPheValMetSerLys 378
Db 1207 CCATCTCTTATTTCTCATGTCAAAA 1233

RESULT 7
ID AAA35311 standard; DNA; 6546 BP.

XX AAA35311;

XX 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:185.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US17712.

XX 03-AUG-1998; 98US-0095212.

PA (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -

XX Disclosure; Page 1331-1333; 1343pp; English.

XX The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1880
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX SQ Sequence 6546 BP; 1603 A; 1692 C; 1500 G; 1750 T; 1 other;

Alignment Scores:

Pred. No.: 4,05e-124 Length: 6546
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: Gaps: 1

US-09-208-629F-3 (1-407) x AAA35311 (1-6546)

QY 11 GlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeuPheLeu 30
Db 127 GGGACTCAGTCAATCAAAATGAAAGCCCTCATCTTTGCAGCTCGCTCGCTCTCTG 186

QY 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 187 TTCCCACTTTTGTGAGAGTGGCATGGAAATGATACAAACACTTGGCAAGCCNACC 246

QY 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 247 TTACCCATTAAAGACCTTCTGTGGAGCTCCCCAAATTCCTTTGAAGAGTTCCCTTTCT 306

QY 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 307 GCCTTGGAAAGCTGGACAGGAGCCAGCATTTACTGTAAAAAATTAAAGTCCCTGAAGAA 366

QY 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 367 GCTTCATCTCCATGTGAAAAATGTACCATGGGTACTGACCATCTCTTAAGTACT 426

QY 110 GlnValIleProAlaIleTyrIleLeuLeuPheValIleGlyValProSerAsnIleVal 129
Db 110 GlnValIleProAlaIleTyrIleLeuLeuPheValIleGlyValProSerAsnIleVal 129

QY 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVal 159
DB 313 TCATCTGTACCACTGATTTCTACCAACTGCGCATTCGCAATTTCTTTTGTGT 372
QY 160 ThrLeuProPheLysIleAlaIleHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
DB 373 ACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTATTGGAGAGGTC 432
QY 180 MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeu 199
DB 433 CTGTGCGGGCCACCAAGTATCTTCTATGGCAACATGACTGCTCCATTCCTCTCT 492
QY 200 ThrCysMetGlyIleAsnArgTyrLeuAlaHisProPheThrTyrGlnLysLeu 219
DB 493 GCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCTTTCACCTACCGGGCGCTG 552
QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
DB 553 CCCAAGCACACCTATGCTTGGTAACATGTGGACTGGTGGGCAACAGTTTCTTATAT 612
QY 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThr 259
DB 613 ATGCTGCCATTTTTCATCTCAAGCAGGAATATATCTTGTTCAGCCAGACATCACCACC 672
QY 260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheVal 279
DB 673 TGCATCATGTTCACCAACTTGGAGTCTCTCATCTCCCTTCCAACTCTATTACTTCATC 732
QY 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThr 299
DB 733 TCCTTGGCATTTCTTGGATTTCTTAATTCATTTGTGCTTATCATCTACTGTCATGCAGCC 792
QY 300 LeuIleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeu 319
DB 793 ATCATCGGACACTTAATGCATACGATCATAGATGGTGTGGTATGTAAGGCGAGTCTC 852
QY 320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis 339
DB 853 CTCATCTCTGGATTTTACATTTGCTTGTCTCCAAAGCAATATATCTTATTTATTCAC 912
QY 340 HisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeu 359
DB 913 CATCTAATCTACTACTACCAACACTGATGGCTTATATTTATATATCTCATAGCTTTG 972
QY 360 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
DB 973 TGCCTGGGTAGTCTTAATAGTGTCTTAGATCCATTCCTTTATTTCTCTCATGTCAAAA 1029

RESULT 9
AAV07373

ID AAV07373 standard; DNA; 1124 BP.

XX AC AAV07373;

XX DT 12-OCT-1998 (first entry)

XX DE Mouse protease-activated receptor 3 (PAR3) genomic DNA.

KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.

XX OS Mus sp.

XX PN W09818456-A1.

XX PD 07-MAY-1998.

XX PF 29-OCT-1997; 97WO-US19732.

XX XX

PR 30-OCT-1996; 96US-0742440.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Connolly A, Coughlin SR, Ishihara H;

XX DR WPI; 1998-271905/24.

XX PT DNA encoding protease-activated receptor 3 - for detection of
XX specific agonists and antagonists, potentially useful for treating
XX e.g. thrombosis, atherosclerosis, inflammation etc.

XX PS Claim 3; Page 39; 74pp; English.

XX CC This sequence comprises genomic DNA for mouse protease-activated
XX receptor 3 (PAR3), a cell surface protein which is specifically
XX activated by thrombin or a thrombin agonist, thereby activating
XX signalling events such as phosphoinositide hydrolysis, calcium ion
XX efflux and platelet aggregation. A cDNA library was established
XX from rat platelets and amplified (primers given in AAV07376-77). A
XX product was identified that was expected to encode a new G-protein
XX coupled receptor, related to PAR1 or 2. This was used to isolate
XX human and murine DNA by a combination of PCR and hybridisation
XX techniques. Murine cDNA (see AAV07372), murine genomic DNA, human
XX cDNA (see AAV07374) and human genomic DNA (see AAV07375) sequences are
XX provided, as well as the murine (see AAW51405) and human (see AAW51406)
XX PAR3 proteins. Also claimed are vectors, host cells and an assay
XX device. Host cells are used to screen compounds for their ability
XX to act as agonists or antagonists of the effects of thrombin-PAR3
XX interaction. Agonists are used to treat wounds, thrombosis,
XX atherosclerosis, restenosis, inflammation and other thrombin
XX activated disorders. Antagonists (see AAW51415-21) are used to
XX control blood coagulation and thereby to treat heart attack and
XX stroke. They also mediate inflammatory and proliferative responses
XX to injury as occur in wound healing, atherosclerosis, restenosis,
XX pulmonary inflammation (ARDS) and glomerulosclerosis.

XX SQ Sequence 1124 BP; 251 A; 314 C; 210 G; 326 T; 23 other;

Alignment Scores:

Pred. No.: 8,64e-112 Length: 1124
Score: 1300.50 Matches: 265
Percent Similarity: 71.99% Conservative: 10
Best Local Similarity: 69.37% Mismatches: 16
Query Match: 60.88% Indels: 92
DB: 19 Gaps: 4

US-09-208-629F-3 (1-407) x AAV07373 (1-1124)

QY 29 PheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAlaLysPro 48

DB 15 TTCCTTTCAATTACA-----GGCATAAATGTTTTCAGAACACTCAGCAAGCCCA 62

QY 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68

DB 63 ACCTTAACCTATTAAAGAGTTTAAATGGGGTCCCAAAATACCTTTCAAGAAATTCNN--- 118

QY 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88

DB 118 ----- 118

QY 89 SerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSer 108

DB 119 NNNNTACACACTCTCCATGTGAATATGTACCATGGGATACCTGAGAGTTCCTTAAGT 178

QY 109 ThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIle 128

DB 179 ACCAAAGTGATACCTGCCATCTACATCCTGCTGTTGTTGATGTTGCTACCGAGCAATC 238

QY 129 valThrIleTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr 148

DB 239 GTGACCCCTGGAAACTCTCCTCAAGACCAATCATCTGTCTGTCTTTCACACC 298

```

QY 149 AsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheIleAlaTyrHi 168
Db 299 AACCTGCCATCGCGATCTCTTTCTGTGTACAGCTGCGTTTAAAGATCNCCTACCA 358
QY 168 sLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrValValPh 188
Db 359 TCTCAATGGCAACAACCTGGGTATTTGGCGAGGTATGTGCGGATCACACGGTCTGTTT 418
QY 188 eTyrGlyAsnMetTyrCysAla---IleLeuIleLeuThrCysMetGlyIleAsnArgTy 207
Db 419 CTACGGCAACATGTACTCGCTANNNTCTGTATCTCACCTGCATGGCATCAACGGCTA 478
QY 207 rLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLe 227
Db 479 CCTGGCCACGGCTCACCTTTTCATACACAGAGCTGCCAAACGCGAGCTTCTCCATGCT 538
QY 227 uMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle-LeuL 247
Db 539 CATGTGTGGCATGGTGTGGTGTATGTTTCTTATACATGCTGCCCTTTGTCTATCCNNA 598
QY 247 ySglnGluTyrHisLeuValHisSerGluIleThrCysHisAspValValAspAlaC 267
Db 599 AGCAGGAGTACCACCTCGTCCATCCGAGATCACCACTGCCAGATGTCTGCGACGGT 658
QY 267 ySglnSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheL 287
Db 659 GCGANTCCCATCATCTCCGATTCCTACTTCTGCTCTCTAGCATTCCTTTGGGTTC 718
QY 287 euIleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerL 307
Db 719 TCATCCGTTTGTGATCATCTCTTGTACAGACTCTCATCCCAAACTTAAATCAA 778
QY 307 ySaspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrI 327
Db 779 AAGATCNGATATGGCTGGGTACATCAAGCGGTCTCTCTCATCTTGTGAATTTCCACCA 838
QY 327 leCysPheAlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyHisA 347
Db 839 TCTGCTTCCCCCCCACCAAG----- 858
QY 347 snThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 367
Db 858 ----- 858
QY 367 ysLeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsnPro**S 387
Db 858 ----- 858
QY 387 erAlaMetAlaArgProLeu***ArgProArgArgAspIleTrpGluAspIleHisAlaT 407
Db 859 -----NNNNNGATATCTGGGAAGACGTACATGCTT 889
QY 407 xp 407
Db 890 GG 891

```

RESULT 10

AAQ84559

ID AAQ84559 standard; cDNA; 2732 BP.

XX

AC AAQ84559;

XX

XX 22-AUG-1995 (first entry)

XX Murine C140 receptor cDNA.

XX

XX G-protein-coupled receptor; G-protein; C140 receptor; ss.

XX Mus musculus.

XX

XX Key Location/Qualifiers

XX 73..1272

XX /*tag= a

FT CDS

FT

XX

PN W09503318-A.

XX

PD 02-FEB-1995.

XX

PF 26-JUL-1994; 94WO-US08536.

XX

PR 26-JUL-1993; 93US-0097938.

XX

PA (CORT-) COR THERAPEUTICS.

XX

PI Scarborough RM, Sundelin J;

XX

DR WPI; 1995-075182/10.

DR

DR P-PSDB; AAR66922.

XX

PT New DNA encoding recombinant C140 receptor - and novel agonists

PT

PT and antagonists and specific antibodies with therapeutic and

PT

PT diagnostic applications.

XX

PS Claim 1; Fig 10; 57pp; English.

XX

CC A cDNA library from a mouse stomach was constructed in lambda gt10
 CC and screened with a probe encompassing the C140 genomic clone (see
 CC AAQ84557). A single phage clone was isolated and cut with EcoRI. The
 CC insert was cloned in pBluescript and pSG5 and sequenced. The
 CC complete nt sequence and deduced AA sequence is given in AAQ84559 &
 CC AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
 CC end. The 5' end of the apparent coding region differs from the 5'
 CC end of the ORF of genomic DNA; it is believed that the 5' end of
 CC the cDNA sequence is correct.

XX

SQ Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;

Alignment Scores:

Pred. No.:	1,71e-44	Length:	2732
Score:	584.00	Matches:	128
Percent Similarity:	51.52%	Conservative:	76
Best Local Similarity:	32.32%	Mismatches:	126
Query Match:	27.34%	Indels:	66
DB:	16	Gaps:	10

US-09-208-629F-3 (1-407) x AAQ84559 (1-2732)

QY 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeu 27

Db 61 CCACGTCGCGGATGCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111

QY 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43

Db 112 ACCCTTCTGGCGGCTCGGTCTCTCTCAGCGGACCGAGAACCTTGACCGGACGCAAC 171

QY 44 AsnSer-----AlaLysProThrLeuThrIleLys 53

Db 172 AACAGTAAAGGAAGAGTCTTATTGCGAGATTAGAACCCAGCCTCAATCCTGGGAAA 231

QY 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGlyGly 73

Db 232 GGGTTCGCGTAGAACCGCTTTTCATCGATGAGTCTCTCGCTCCATCTCACCACGGG 291

QY 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93

Db 291 ----- 291

QY 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113

Db 292 -----AAGCTGACCACGGTCTTTCTTCG 315

QY 114 AlaIleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTrpLys 133

Db 316 GTGCTACATTATTGTGTTTGATGTTTCCAGTAATGCGATGCGCCCTCTGATC 375

QY 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152

```

Db      376  TTCTTTTCGACGAGAAAGAACACCCCGCGGTGATTTACATGGCCACCTGGCTTG 435
Qy      153  AlaAspLeuLeuPheCysValThrLeuProPheLeuAlaTyrHisLeuAsnGlyAsn 172
Db      436  GCGACCTCTCTCTGTCATCTGGTTCCTCCCGCTGAAGATCTCTACCACTACATGGCAAC 495
Qy      173  AsnTyrValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db      496  AACTGGGTCTACGGGGAGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTATACATG 555
Qy      193  TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Db      556  TATTGCTCCATCTCTTCATGACCTGCTCAGCGTGCGAGGTACTGGGTGATCGTGAAC 615
Qy      213  ProPheThrTyrGlnLeuLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db      616  CCCATGGGACACCCCGAGG---AAGAGGCAACATCCCGTTCCTGCGTCTCTTTGGCAATC 672
Qy      233  TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
Db      673  TGGCTCTGATTTTCTGGTCACCATCCCTTTGTATGTCATGAGCAGACCATCTACAT 732
Qy      253  ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db      733  CCAGCATTTGAACATCACCATCTGTCCAGATGTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy      273  PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db      790  GACATGTTCAATTAATCTCTCTCACTGCGCATTTGGAGTCTTCTGTTCCCGGCCCTCT 849
Qy      293  IleIlePheCysTyrThrThrLeuLeu-----ATCCGACTCATCATCACCGTGTG 957
Db      850  ACTGCATCTGCTACGTGCTCATGATCAAGACGCTCGCTCTCTGCTGATGGATGAACAC 909
Qy      303  LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db      910  TCAGAGAAGAAAGGACAGAGGCT-----ATCCGACTCATCATCACCGTGTG 957
Qy      323  ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db      958  GCCATGTAATCTATCTGCTGTTGCTTCTAGCAACCTTCTGCTGAGTGCATTTATTTCT 1017
Qy      343  TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db      1018  ATCAAAACCCAGGAGGAGCCAGCTGTACGCCCTCTACCTTGTGCGCCCTCTGCTGCTG 1077
Qy      363  SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db      1078  ACCCTCAACAGCTCATAGACCCCTTTGCTCTATTACTTTGTCTCAAAA 1125

RESULT 11
AAT32038
ID   AAT32038 standard; cDNA; 2732 BP.
AC   AAT32038;
XX
XX
DT   02-APR-1997 (first entry)
DE
DE   Murine C140 receptor cDNA clone.
KW   C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW   antagonist; hypertension; hypotension; blood pressure; ss.
XX
OS   Mus sp.
XX
FH   Key
FT   CDS
FT   Location/Qualifiers
FT   73..1272
FT   /tag= a
FT   sig_peptide
FT   73..165
FT   /tag= b
FT   /note= "the signal sequence differs from a
FT   signal sequence shown in a genomic DNA

```

```

FT      mat_peptide 166..1269
FT                                     /*tag= c
XX
XX  WO9623225-A1.
XX
XX  01-AUG-1996.
XX
XX  25-JAN-1996; 96WO-0501179.
XX
XX  25-JAN-1995; 95US-0390301.
XX
XX  (CORT-) COR THERAPEUTICS INC.
XX
XX  Scarborough RM, Sundelin J;
XX
XX  WPI; 1996-362813/36.
XX
XX  P-PSDB; AAW01954.
XX
XX  Vector for expression C140 cell surface receptor in host cell
XX  useful to identify C140 agonist and antagonists, which are
XX  antihypertensives and elevators of blood pressure, respectively
XX
XX  Example 4; Fig 10A-B; 60pp; English.
XX
XX  AAT32038 encodes the murine C140 receptor (C140R). The sequence may be
XX  engineered so as to allow the recombinant expression of C140R in a
XX  suitable host cell, i.e. by removing the native expression-control
XX  sequence and replacing them with control sequences operable in the
XX  host. Such a recombinant receptor can be expressed on the surface of
XX  oocytes, this provides a good assay system for identifying
XX  agonists/antagonists of C140R. The C140 receptor is a G-protein linked
XX  receptor and a member of the "seven-pass" transmembrane receptor
XX  superfamily (peptide chain of the receptor passes through the cell
XX  membrane seven times, producing seven transmembrane regions within
XX  the receptor molecule). The C140 receptor is involved in controlling
XX  blood pressure. C140 antagonists (see AAW01942-W01951) are useful to
XX  inhibit signalling from this receptor, resulting in an increase in
XX  blood pressure and are therefore useful in pharmaceuticals for the
XX  treatment of hypotension (low blood pressure). Conversely agonists
XX  (see AAW01914-W01941) of C140 are useful in pharmaceuticals for the
XX  treatment of hypertension (high blood pressure).
XX
XX  Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 1.71e-44 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 17 Gaps: 10

US-09-208-629F-3 (1-407) x AAT32038 (1-2732)

Qy      8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db      61 CCACGTCGCGGGATGCGAAGTCTTCAGCCTGGCGTGGCTGCTG-----GGAGGTATC 111
Qy      28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db      112 ACCCTTCTGGCGCCCTCGGTCTCTCTGACGCCGACCGAGAACCTTCGACCGGACGCAAC 171
Qy      44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db      172 AACAGTAAAGGAAGAAGTCTTATTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAA 231
Qy      54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db      232 GGGGTTCCGGTAGAACACGAGCTTTTCCATCGATGATGAGTCTTCTCGGTCCATCTCCACCGGG 291

```

segment encoding this receptor (see AAT32036),
the signal sequence given for this clone is
believed to be the correct sequence"


```
Qy 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTTCATGACCTGCCTCAGCGTGAGAGGTACTGGGTGATCGTGAACCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCAGG--AAGAAGGCAACATCGCGTGGCGTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 823 CTCCTGATTTTCGGTCACATCCCTTTGTATGTCAAGAGCAGACCATCTACATCCA 882
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 883 GCATTGAACATCACCATCTGCAGATGTCTGCTGAG--GAGGTATTGGTGGGGAC 939
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPheValIle 293
Db 940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCGGGCCCTCTTACT 999
Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
Db 1000 GCATCTCCCTACGCTCATCATCAAGACGCTCGCTCTTCTGCTATGGATGAACACTCA 1059
Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
Db 1060 GAGAACAAAGGAGAGGGCT-----ATCGACTCATCATCAGCGTGTGGCC 1107
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
Db 1108 ATGTAATTCTCTCTTCTGCTCTAGCAACCTTCTGCTGATGCAATTATTTCTTAATC 1167
Qy 344 TyrTyrHisThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGlySer 363
Db 1168 AAAACCCAGGAGGACGACGCTACGCTTACCTTGTGCGCCCTCTGCTGCGAC 1227
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1228 CTCACAGCTGCATAGACCCCTTGTCTATTAATCTTCTCTCAAA 1272

RESULT 13
AAT32036
ID AAT32036 standard; DNA; 1477 BP.
XX
AC AAT32036;
XX
DT 01-APR-1997 (first entry)
XX
DE Murine C140 receptor-encoding DNA.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 232..1419
FT sig_peptide /*tag= a
FT 232..312
FT /*tag= b
FT /*note= "putative signal sequence, differs from
FT signal sequence shown in a cDNA clone of
FT this receptor (see AAT32038), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT mat_peptide 313..1416
FT /*tag= c
XX
PN WO9623225-A1.
XX
PD 01-AUG-1996.
XX
PF 25-JAN-1996; 96WO-US01179.
```

```
XX
PR 25-JAN-1995; 95US-0390301.
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Scarbrough RM, Sundelin J;
XX WPI; 1996-362813/36.
XX P-PSDB; AAW01952.
XX
XX Vector for expression C140 cell surface receptor in host cell
XX useful to identify C140 agonist and antagonists, which are
XX antihypertensives and elevators of blood pressure, respectively
XX
XX Example 1; Fig 1A-B; 60pp; English.
XX
XX AAT32036 encodes the murine C140 receptor (C140R), including a tentative
XX signal sequence. The sequence may be engineered so as to allow the
XX recombinant expression of C140R in a suitable host cell, i.e. by
XX removing the native expression-control sequence and replacing them with
XX control sequences operable in the host. Such a recombinant receptor
XX can be expressed on the surface of oocytes, this provides a good assay
XX system for identifying agonists/antagonists of C140R. The C140
XX receptor is a G-protein linked receptor and a member of the "seven-
XX pass" transmembrane receptor superfamily (peptide chain of the
XX receptor passes through the cell membrane seven times, producing
XX seven transmembrane regions within the receptor molecule). The C140
XX receptor is involved in controlling blood pressure. C140 antagonists
XX (see AAW01942-W01951) are useful to inhibit signalling from this
XX receptor, resulting in an increase in blood pressure and are therefore
XX useful in pharmaceuticals for the treatment of hypotension (low blood
XX pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
XX in pharmaceuticals for the treatment of hypertension (high blood
XX pressure).
XX
SQ Sequence 1477 BP; 341 A; 415 C; 319 G; 402 T; 0 other;

Alignment Scores:
Pred. No.: 9, 03e-44 Length: 1477
Score: 572.50 Matches: 110
Percent Similarity: 59.32% Conservatives: 65
Best Local Similarity: 37.29% Mismatches: 103
Query Match: 26.80% Indels: 17
DB: 17 Gaps: 5

US-09-208-629F-3 (1-407) x AAT32036 (1-1477)
Qy 95 ValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTCTCTCGTCCATCTCACCGGAAGCTGACACCGGTCTTCTTCGGTC 465
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 525
Qy 135 SerLeuArgThr--LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAACGAGAGAAACACCCCGCGTGATTTACATGGCCACCTTGGCTTGGCC 585
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTCTCATCTGTTTCCCTCCCTGAGAGATCTCTTACCACCTACATGCAACAC 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGAT 705
Qy 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTTCTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
```


QY 301 ILeHisLysLeuYsSer-----LysAspArgIleTrpLeuGlyTyr 314
 DB 854 ATCAGATGTCGGATCTTTCGCCATGGATGAAACTCAGAGAGAAAGAGAGGGCC 913
 QY 315 ILeYsAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 DB 914 ATCAAACTCATGTGTCACCTGCTGGCCATGTACCTGATCTGCTTCACTCTAGTAACCTT 973
 QY 335 ILeLeuValIleHisAlaAsnTyrTyrTrpHisAsnThrAspSerLeuTyrPheMet 354
 DB 974 CTGCTTGTGGTCATTAATTTCTGATTAGAGCCAGGCCAGCCATGCTATGCCCTG 1033
 QY 355 TyrLeuIleAlaLeuLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
 DB 1034 TACATTGAGCCCTCTGCCCTCTTACCCTTAACAGCTGCATCGACCCCTTGTCTATTAC 1093
 QY 375 ValMetSer 377
 DB 1094 TTTGTTTCA 1102
 RESULT 15
 AAT32037
 ID AAT32037 standard; DNA; 1255 BP.
 XX
 AC AAT32037;
 XX
 DT 01-APR-1997 (first entry)
 XX
 DE Human C140 receptor-encoding DNA.
 XX
 KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
 KW antagonist; hypertension; hypotension; blood pressure; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 56..1197
 FT /*tag= a
 FT sig_peptide 56..136
 FT /*tag= b
 FT /note= "putative signal sequence, differs from
 FT signal sequence shown in a cDNA clone of
 FT this receptor (see AAT32039 - this clone also
 FT encodes a mature receptor having four amino
 FT acid changes, when compared to this sequence)"
 FT mat_peptide 137..1194
 FT /*tag= c
 XX
 FN WO9623225-A1.
 PD 01-AUG-1996.
 XX
 PF 25-JAN-1996; 96WO-US01179.
 XX
 PR 25-JAN-1995; 95US-0390301.
 XX
 PA (CORT-) COR THERAPEUTICS INC.
 XX
 PI Scarborough RM, Sundelin J;
 DR WPI; 1996-362813/36.
 DR P-PSDB; AAW01953.
 XX
 FT Vector for expression C140 cell surface receptor in host cell
 FT useful to identify C140 agonist and antagonists, which are
 FT antihypertensives and elevators of blood pressure, respectively
 XX
 PS Example 2; Fig 2A-B; 60pp; English.
 CC
 CC AAT32037 encodes the human C140 receptor (C140R), including a tentative
 CC signal sequence. The sequence may be engineered so as to allow the
 CC recombinant expression of C140R in a suitable host cell, i.e. by

removing the native expression-control sequence and replacing them with
 control sequences operable in the host. Such a recombinant receptor
 can be expressed on the surface of oocytes, this provides a good assay
 system for identifying agonists/antagonists of C140R. The C140
 receptor is a G-protein linked receptor and a member of the "seven-
 pass" transmembrane receptor superfamily (peptide chain of the
 receptor passes through the cell membrane seven times, producing
 seven transmembrane regions within the receptor molecule). The C140
 receptor is involved in controlling blood pressure. C140 antagonists
 (see AAW01942-W01951) are useful to inhibit signalling from this
 receptor, resulting in an increase in blood pressure and are therefore
 useful in pharmaceuticals for the treatment of hypotension (low blood
 pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
 in pharmaceuticals for the treatment of hypertension (high blood
 pressure).
 XX
 SQ Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 other;
 Alignment Scores:
 Pred. No.: 5,61e-43 Length: 1255
 Score: 563.00 Matches: 115
 Percent Similarity: 54.52% Conservative: 72
 Best Local Similarity: 33.53% Mismatches: 128
 Query Match: 26.36% Indels: 28
 DB: 17 Gaps: 7
 US-09-208-629f-3 (1-407) x AAT32037 (1-1255)
 QY 49 ThrLeuThrIleYsSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
 DB 116 ACCCTTGTCTTCTTCTTGTACAGGACCAATAGATCTCTAAAGGAAGACCTTATT 175
 QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
 DB 176 GGTAAGGTGATGGGCACATCCACGTCCTCAAGAAAGAGGTACAGTTGAA----- 226
 QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
 DB 227 -----ACAGTCITTTCTGTGGATGAGTTTCTGTCATCTGCTCTCACTGGA 271
 QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
 DB 272 AAACCTGACCACCTGCTTCTCTCCAAATGTCTACAAATGTGTGTGTGGTGGTGGTGGCA 331
 QY 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
 DB 332 AGTAACGGCATGGCCCTGTGGGTCTTTCTTTCCGAACCTAAGAAAGAACCCCTGCTGTG 391
 QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
 DB 392 ATTTACATGGCCAATCTGGCTTGGCTGACCTCTCTCTCATCTGGTTCCCTTGAAG 451
 QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 DB 452 ATTGCCTATCACAATACATACGCAACAACTGGAATTTATGGGAAGCTCTTTGTATATGCTT 511
 QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIle 204
 DB 512 ATGGCTTTTCTATGGCAACATGTAATCTTCTTCAATCTCTTCAATGACCTGCTCAGTGTG 571
 QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 DB 572 CAGAGGTATGGGTATCGTGAACCCCATGGGSCACTCCAGG---AAGAAGGCAACATT 628
 QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPheVal 244
 DB 629 GCCATTGGCATCTCCCTGGCAATATGCTGCTGATTTCTGCTGCACCATCCCTTGTAT 688
 QY 245 IleLeuLysGlnGluThrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 DB 689 GTGCTGAAGCAACACCATCTTCTGCTGCTGCAACATCAGACCTGTCTAATGATTTTG 748
 QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280

Db 749 -----CTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 793
 Qy 281 LeuAlaPheGlyLeuLeuProPheValIleIlePheCysTyrThrThrLeu 300
 Db 794 CTGGCCATTGGGCTTTCTGTTCCAGCCTTCTCTACAGCCTCTGCTATGCTGATG 853
 Qy 301 IleHisTysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
 Db 854 ATCAGAATGCTGCATCTTCTGCCATGATCAAACTCAGAGAAGAAAGAGAGGCC 913
 Qy 315 IleLysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 914 ATCAAACTATGTGCTGCTGCGCATGTACCTGATCTGCTTCACTCTAGTAACCT 973
 Qy 335 IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
 Db 974 CTGCTCTGGTGCAATTATTTCTGATTAAAGGCCAGGCCGATGCTATGCCCTG 1033
 Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPhe 374
 Db 1034 TACATTGAGCCCTCTGCTCTCTTACCCTTAACAGCTGCATCGACCCCTTTGCTATTAC 1093
 Qy 375 ValMetSer.377
 Db 1094 TTGTTTCA 1102

RESULT 16

ID AAZ50773 standard; DNA; 1289 BP.

AC AAZ50773;

DT 31-MAY-2000 (first entry)

DE Human protease activated receptor-2 DNA.

EX Human; PAR-2; protease activated receptor;

KW antisenese molecule; PAR antibody; cytostatic; therapeutic;

KW metastatic tumour cell; placental implantation; invasive cell; ds.

OS Homo sapiens.

PN WO200008150-A1.

PD 17-FEB-2000.

PF 05-FEB-1999; 99WO-IL00079.

PR 07-AUG-1998; 98IL-0125698.

PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

PI Bar-Shavit R;

DR WPI; 2000-205706/18.

PT Treating metastatic tumor cells useful for treating disorders involving placenta implantation in a female comprises administration of an antisenese molecule complementary to an RNA sequence of a protease activated receptor protein -

PS Example 3; Fig 9; 46pp; English.

CC The patent discloses a method to treat metastatic tumour cells using an antisenese molecule comprising a polynucleotide complementary to an RNA sequence of a protease activated receptor (PAR) protein, or an antibody capable of binding to a PAR protein. The antisenese molecules and antibodies of PAR protein are also used to treat disorders associated with implantation of placenta. The present sequence is a human PAR-2 DNA used for producing antisenese molecules for treating invasive cells.

CC XX

SQ Sequence 1289 BP; 300 A; 320 C; 270 G; 399 T; 0 other;

Alignment Scores:

Pred. No.: 5-81e-43 Length: 1289
 Score: 563.00 Matches: 115
 Percent Similarity: 54.52% Conservative: 72
 Best Local Similarity: 33.53% Mismatches: 128
 Query Match: 26.36% Indels: 28
 DB: 21 Gaps: 7

US-09-208-629f-3 (1-407) x AAZ50773 (1-1289)

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
 Db 72 ACCCTTGTCTTCTTCTTGTACAGAACCAATAGATCTCTTAAGGAAGAGCCCTTATT 131
 Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
 Db 132 GGTAAGGTGTAGTGGCAGCATCCACGTCACCTGGAAGAGGAGTTACAGTTGAA----- 182
 Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
 Db 183 -----ACAGTCTTTCTGTGGATGAGTTTCTGCATCTGTCTCTCACTGGA 227
 Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
 Db 228 AAACAGCACCCTGTCTTCTTCCAAATTGTCTACAAATTGTGTGTGGTGGTGTGGCCA 287
 Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr-----LysSerIleSerLeuVal 144
 Db 288 AGTAACGGCATGCCCTGTGGGTCTTTCTTTCCGAACCTAAAGAAAGAACCCCTGTGTG 347
 Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
 Db 348 ATTACATGGCCCAATCTGGCTTGGTGGCTGACCTCTCTGTCTCATCTGGTTCCTTCAAG 407
 Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 Db 408 ATTGCTATCACAATACATACATGCAACAACTGATTTATGGGAAGCTCTTTGTAATGCTT 467
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
 Db 468 ATTGGCTTTTCTATGGCAACATGATCTGTTCATTTCTTCATGACCTGCCTCAGTGTG 527
 Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 528 CAGAGGTATGGGTCTATCGTGAACCCCATGGGCACCTCCAGG---AAGAAGGCAACATT 584
 Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
 Db 585 GCCATTGGCATCTCCCTGGCAATATGGTCTGATTTCTGTGTGCACCATCCCTTTGTAT 644
 Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 Db 645 GTCTGAAGCAGACCATCTTCATTCCTGCCCTGCAACATCACGACCTGTGATGTTGTG 704
 Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
 Db 705 -----CCTGAGCAGCTTGTGGGGAGACATGTTCAATTACTTCTCTCT 749
 Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 750 CTGGCCATTGGGTCTTTCTGTTCCTCCAGCCTTCTCCACAGCCTCTGCCTATGTGTGATG 809
 Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
 Db 810 ATCAGAATGCTGCGATCTTCTGCCATGATGAACCTCAGAGAAGAAAGAGAGAGGCC 869
 Qy 315 IleLysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 870 ATCAAACTCATGTGCTACTGTCTGGCCATGTACCTGATCTGCTTCACTCTCTAGTACCTT 929
 Qy 335 IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

Df

Qy

251 GTCCCTCGGTGGAAAACTGACCAACTGTCCTTCCCTTTCCAATATTGCTTACAACAATAATTGTTTTGCCG 310
122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
|||||:::|||||:::|||:::|||:::|||:::|||
311 GTGGGTTTGGAAGTAACGCATGGCCTATTGGTCTTTCTTCCGAACTAAGAAGAAG 370
141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuCysValThr 160
|||||:::|||||:::|||:::|||:::|||:::|||
371 CACCCTCGCTGATTACATGGCAAATCTGGCCTTGCGTAGCACCTCTCTCTCATCTGG 430
161 LeuProPheLysIleAlaTyrrHisLeuAsnGlyAsnAsnTrpValPheGlyGlutMet 180
|||||:::|||||:::|||:::|||:::|||:::|||
431 TTCCCCTTGAAGATTGCCATTACATACATACATGGCAACACTGGATTTATGGGAAGCTCTT 490
181 CysArgIleThrThrValValPheTyrrGlyAsnMetTyrrCysAlaIleLeuLeuThr 200
|||||:::|||||:::|||:::|||:::|||:::|||
491 TGTAAATGCTTATTGGCTTTTCTATCGCAACATGFACTGTTCCATTCTCTTCATGACC 550
201 CysMetGlyIleAsnArgTyrrLeuAlaThrAlaHisProPheThrTyrrGlnLysLeuPro 220
|||||:::|||||:::|||:::|||:::|||:::|||
551 TGCCCTCAGTGTGAGAGGATTGGGTGCATGCTGAACCCCATCGGGCACTCCAGG--AAG 607
221 LysArgSer:PheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrrMet 240
|||||:::|||||:::|||:::|||:::|||:::|||
608 AAGCGAAAATTGCCAATTGGCACTCCCTGGCAATATGGCTGTCTGACTCTGCTGGTCAACC 667
241 LeuProPheValIleLeuLysGlnTuTyrrHisLeuValHisSerGluIleThrThrCys 260
|||||:::|||||:::|||:::|||:::|||:::|||
668 ATCCCTTTGTATGCTGTGAAGCAGACCACTTCTTCCTGCCCTGAAACATCAGCAGCTGT 727
261 HisAspValValAspAlaCysGluSerProSerSer:Phe-----Ar:gPheTyrr 276
|||||:::|||||:::|||:::|||
728 CATGATGTTTG- - - - - CCTGAGCAGCTCTTGGTGGGAGACATGTTTCAAT 772
277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIlellellepHeCys 296
|||||:::|||||:::|||:::|||:::|||
773 TACTTCCTCTCTCTGCCAATTGGGGTCTTCTGTTCCCCAGCCTTCTCAGAGCCTCTGCC 832
297 TyrrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
|||||:::|||||:::|||:::|||
833 TAGTGCTGATGATCAGATGCTGGCATCTTCTGCCATGATGAAACTCAGAGAAGAAA 892
311 TrpLeuGlyTyrrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAla 330
|||||:::|||||:::|||:::|||:::|||
893 AGGAAGAGGGCCATCAAACTCATTTGTCCTGGCATGTACCTGATCTGCTTCACT 952
331 ProThrAsnIlelleLeuValIleHisHisAlaAsnTyrrTyrrHisAnthrAppSer 350
|||||:::|||||:::|||:::|||
953 CCTAGTAACCTCTGCTGTGGTGCAATATTTCTGATTAAAGCCAGGCCAGAGCCAT 1012
351 LeutyrrPheMetTyrrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
|||||:::|||||:::|||:::|||
1013 GTCATGCCCCTGACATTGTAGCCCTCTGCCTCTCTACCCCTAACAGCTGCATCGACCCC 1072
371 PheLeutyrrPheValMetSer 377
|||||:::|||||:::|||
1073 TTGTGCTATTACTTCTGTTTCA 1093

RESULT 19
AAT32039
ID AAT32039 standard; cDNA; 1414 BP.
XX
X X AAT32039;
X X DT
X X 02-APR-1997 (first entry)
X X Human C140 receptor cdna clone.
X X C140 receptor; G-protein linked; coupled; seven pass; agonist;
X W antagonist; hypertension; hypotension; blood pressure; ss.
X X Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 50..1243
 FT /*tag= a
 FT sig_peptide 50..130
 FT /*tag= b
 FT /note= "the signal sequence, differs from the
 FT signal sequence shown in a genomic clone of
 FT this receptor (see AAT32037). The cDNA clone
 FT also encodes a mature receptor having four
 FT amino acid changes, when compared to the
 FT genomic clone"
 FT mat_peptide 131..1240
 FT /*tag= c
 XX W05623225-A1.
 PD 01-AUG-1996.
 XX 25-JAN-1996; 96WO-US01179.
 XX 25-JAN-1995; 95US-0390301.
 XX (CORT-) COR THERAPEUTICS INC.
 XX Scarbrough RM, Sundelin J;
 XX WPI; 1996-362813/36.
 DR P-PSDB; AAW01955.
 XX Vector for expression C140 cell surface receptor in host cell -
 FT useful to identify C140 agonist and antagonists, which are
 FT antihypertensives and elevators of blood pressure, respectively
 XX Example 5; Fig 11A-B; 60pp; English.
 XX AAT32039 encodes the human C140 receptor (C140R). The sequence may be
 CC engineered so as to allow the recombinant expression of C140R in a
 CC suitable host cell, i.e. by removing the native expression-control
 CC sequence and replacing them with control sequences operable in the
 CC host. Such a recombinant receptor can be expressed on the surface of
 CC oocytes, this provides a good assay system for identifying
 CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
 CC receptor and a member of the "seven-pass" transmembrane receptor
 CC superfamily (peptide chain of the receptor passes through the cell
 CC membrane seven times, producing seven transmembrane regions within the
 CC receptor molecule). The C140 receptor is involved in controlling blood
 CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
 CC signalling from this receptor, resulting in an increase in blood
 CC pressure and are therefore useful in pharmaceuticals for the treatment
 CC of hypertension (low blood pressure). Conversely agonists (see AAW01914-
 CC AAW01941) of C140 are useful in pharmaceuticals for the treatment of
 CC hypertension (high blood pressure).
 XX SQ Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 other;
 Alignment Scores:
 Pred. No.: 8,61e-42 Length: 1414
 Score: 551.00 Matches: 118
 Percent Similarity: 51.77% Conservative: 72
 Best Local Similarity: 32.15% Mismatches: 139
 Query Match: 25.80% Indels: 38
 DB: 17 Gaps: 8
 US-09-208-629F-3 (1-407) x AAT32039 (1-1414)
 Qy 22 LeuValAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 Db 74 CTGCTGGGGCGCCACTCTCTCTAGCAGCCTCTCTCTCTGCTGAGCCACCATCCAAGGA 133
 Qy 42 SerAspAsnSerAlaGlyLeuLeuPheLeuThrLeuThrIleLeuSerPheAsnGlyProGlnAsn 61
 Db 134 ACCAATAGATCCCTTAAGAAGGAAGACGCTTATTGTAAGGTTGATGCGC----- 181

QY 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIle 81
 DB 182 -----ACATCCCAAGTTCAGTT 214
 QY 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
 DB 215 GAA-----ACAGTCTTTCTGTGTGATGAGTTTCTGCATCT 250
 QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleIleThrIleLeuLeuPheVal 121
 DB 251 GTCTCGCTGGGAAACTGACCACTGCTCTCTCCATTCCTACCAATGTCTACCAATGTGTGTTGG 310
 QY 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
 DB 311 GTGGGTTTGCCTCAAGTAACGGCATGCGCTTATGGTCTTTCTTCGGAACCTAAGAAGAAG 370
 QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 DB 371 CACCCCTGCTGTGATTTACATGGCCAACTCTGGCTTGGCTGACCTCTCTCTGTCATCTGG 430
 QY 161 LeuProPheLeuIleAlaIleThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 DB 431 TTCCCTTGAAGATTGCTTATCACATACATGCACTGCACTGTTCCATCTCTTCATGACC 550
 QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
 DB 491 TGTAAATGCTTATTGGCTTTTCTATCGCAACATGACTGTTCCATCTCTTCATGACC 550
 QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 DB 551 TGCTCAGTGTGCAGAGGATTGGGTGTCATCGTGAACCCCATGGGCACCTCCAGG---AAG 607
 QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 DB 608 AAGGCAAACTTGGCATCTCCCTGGCAATATGCTGCTGACTCTGTGTGTCACC 667
 QY 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
 DB 668 ATCCCTTTGATGTCGTGAAGCAGACCATCTTCTCTCTGCTGCAACATCAGACCTGT 727
 QY 261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
 DB 728 CATGATGTTTG-----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772
 QY 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCys 296
 DB 773 TACTTCTCTCTGTCGCAATGGGGTCTTTCGTCCAGCCTTCTCAGACCTCTGCC 832
 QY 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
 DB 833 TATGCTGATGATCAGATGCTGCGATCTTCTGCCATGGATGATAAACTCAGAGAAGAAA 892
 QY 311 TrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAla 330
 DB 893 AGGAAGAGGGCCATCAAACTCATGTGCTGCTGGCATGTACCTGATCTGCTTCACT 952
 QY 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
 DB 953 CCTAGTAACCTTCTGCTTGTGGTCATTAATTTCTGATTAAGACCGAGGCCAGACCAT 1012
 QY 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
 DB 1013 GTCTATGCCCTGTACATTGTAGCCCTCTGCTCTCTACCTTAACAGCTGCATCGACCCC 1072
 QY 371 PheLeuTyrPheValMetSer 377
 DB 1073 TTGTCTATTACTTTGTTCA 1093
 RESULT 20
 AAT93367
 ID AAT93367 standard; cDNA; 1312 BP.
 XX
 AC AAT93367;

XX 01-MAY-1998 (first entry)
 XX Mouse thrombin receptor cDNA.
 XX Transgenic animal; thrombin receptor; animal model; restenosis;
 KW knockout animal; antagonist; agonist; mouse; murine;
 KW Alzheimer's disease; wound repair; ss.
 XX Mus sp.
 XX Key Location/Qualifiers
 PH complement (612..631)
 FT /*tag= a
 FT /label= MTR-3
 FT 1294..1311
 FT /*tag= b
 FT /label= MTR-2
 XX W09740138-A1.
 XX 30-OCT-1997. 97WO-US06575.
 XX 22-APR-1997; 97WO-US06575.
 XX 25-JUN-1996; 96US-0020544.
 XX 23-APR-1996; 96US-0016070.
 XX (ORTH) ORTHO PHARM CORP.
 XX Andrade-Gordon P, Leung W, Zhou L;
 XX WPI; 1997-535823/49.
 XX Transgenic animals with altered form of thrombin receptor gene -
 PT useful in study of receptor function and identification of
 PT antagonists and agonists for disease therapy
 XX Example 1; Fig 1; 47pp; English.
 CC The present sequence was used in the preparation of a transgenic
 CC non-human vertebrate with somatic and germ cells containing an
 CC altered form of a thrombin receptor (TR) gene, where the altered
 CC gene has replaced a wild type TR gene in the animal or an ancestor
 CC at an embryonic stage using embryonic stem cells. The transgenic
 CC animal provides an animal model to understand receptor function in,
 CC e.g. platelets, endothelial cells and neurons, and evaluate drug
 CC therapies modulating TR function/expression in human cells.
 CC Alterations rendering the wild type gene non-functional produce
 CC knockout animals, useful as non-human whole animal models for
 CC diseases involving TR equivalents in humans. Alternatively,
 CC substituting the naturally occurring gene for a gene from a 2nd
 CC species or with a mutation, results in animals producing the 2nd
 CC species or mutated gene products. These transgenic animals are
 CC useful for drug antagonist and agonist studies, creation of animal
 CC models of human diseases and investigation of disorders associated
 CC with human TR-mediated responses. Antagonists may be used
 CC therapeutically for cardiovascular problems, e.g. restenosis, to
 CC limit platelet adhesion during interventional procedures or to
 CC generate therapies for bone degenerative and neurodegenerative
 CC diseases, e.g. Alzheimer's. Agonists may be useful in wound repair.
 XX SQ Sequence 1312 BP; 247 A; 406 C; 328 G; 331 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.97e-38 Length: 1312
 Score: 514.50 Matches: 120
 Percent Similarity: 51.28% Conservative: 80
 Best Local Similarity: 30.77% Mismatches: 143
 Query Match: 24.09% Indels: 47
 DB: 18 Gaps: 13
 US-09-208-629F-3 (1-407) x AAT93367 (1-1312)

QY 20 LeuLeuValAlaAlaGly-----LeuLeuPheLeuProValThr 33
 |||:|||||
 Db 30 TTGCTGATCGTCGCCCTCGCCTCGAGCTGTGGGTCCCTTGCTGCTTCCCGCGTCCCT 89
 |||:|||||
 QY 34 ValCysGlnSerGlyIleAsnValSerAspAsnSerAlaIysProThrLeuThrIleLys 53
 |||:|||||
 Db 90 ATGAGCCAGCCAGAAATCAGAGAGGAGAGATGCTACGGTGAACCCC-----CGC 137
 |||:|||||
 QY 54 SerPhe-----AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
 |||:|||||
 Db 138 TCATTCTTTCTTAAGGAATCCAGTGAAATATACATTTGAACCTGGTCCCTTGGGGGATGAG 197
 |||:|||||
 QY 72 GluGly-----TrpThrGlyAlaThrThrThrIleLysAla 83
 |||:|||||
 Db 198 GAGGAGGAGGAGAAAAATGAAAGCGTCTCGTGGAGGGTGGGCGAGTCTACTTAAATATA 257
 |||:|||||
 QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
 |||:|||||
 Db 258 AGCTCCCTCCTCACACGCGCTCTCCCTTCATCTCCGAGGAGCGCTCCGATATCTG 317
 |||:|||||
 QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
 |||:|||||
 Db 318 ACCAGCCCTGGCTGAGCTCTTCATGCCCTCCGTGTACAGATGTGTTCAITGTGACG 377
 |||:|||||
 QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys---SerIleSer 142
 |||:|||||
 Db 378 CTTCTCTGAACGTCCTCGCCATCGCAGTGTTCGTCTTGAGGATGAAGGTCAAGAAGCCG 437
 |||:|||||
 QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
 |||:|||||
 Db 438 GCGGTGTGTACATGCTGCACCTGGCCATGGCGACGTGCTCTTCGTGCGGTGCTCCCT 497
 |||:|||||
 QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
 |||:|||||
 Db 498 TCACAGATCAGCTACTACTTCCGGCACTGATTTGGCAGTTTCGGGTCTGGAATGTGCCGT 557
 |||:|||||
 QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMet 202
 |||:|||||
 Db 558 TTGCGCACCGCAGCGTTTACTATACATGTAGCCCTCCATCATGTCATGACGGTCATA 617
 |||:|||||
 QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeu 219
 |||:|||||
 Db 618 AGCATTGACGGTTCCTGGCGGTGGTATCCGATCCAGTCCCTCCCTGGCGCACTCTG 677
 |||:|||||
 QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
 |||:|||||
 Db 678 GCGGAGCGCAACTTCACT-----TGCGTGTTCATTTGGGTGATGGCCATCATGGGG 728
 |||:|||||
 QY 240 MetLeuProPheValIleLeuLysGlnIuTyrHisLeuValHisSerGluIleThrThr 259
 |||:|||||
 Db 729 GTGGTCCCTTCTCTCAAGGAGCAGACCCCGAGTTCGGGACTCAACATCACCACC 788
 |||:|||||
 QY 260 CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheVal 279
 |||:|||||
 Db 789 TGGCAGCATCTCTCAGTGAG---AACCTGATGCAAGGCTTTTACTCGTACTACTTCTCG 845
 |||:|||||
 QY 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThr 299
 |||:|||||
 Db 846 GCCTTCTCCGCATCTTCTTTCTTGCGCGTGTGATCGTTCCACGGTCTGTCACAGTCC 905
 |||:|||||
 QY 300 LeuIleHisLysLeuLysSer-----LysAspArgIleTrp 311
 |||:|||||
 Db 906 ATCATCCGGTGTCTGAGCTCTCCGGTTCGCCAGCGGCAAGAAGTCGCGGCTTTCG 965
 |||:|||||
 QY 312 LeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 331
 |||:|||||
 Db 966 -----TTCTGTCTGCGCG-----GTGTTCTGCATCTTTCATCTGCTGCTTGGGCC 1013
 |||:|||||
 QY 332 ThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyr-----HisAsp 347
 |||:|||||
 Db 1014 ACCACGTCCTCTGATTGTGCAC-----TACCTTTTCTCTCCGACAGTCTCTGGT 1064
 |||:|||||

ID	AAZ32191	standard; cDNA; 3472 BP.
XX	AAZ32191;	
AC	13-JAN-2000	(first entry)
XX		
DT		
XX		
DE	Human thrombin receptor nucleotide sequence.	
XX		
KW	Human; coding sequence polymorphism; vascular pathology gene;	
KW	polymorphic site; phenotype correlation; forensic; paternity testing;	
KW	medicine; genetic analysis; vascular disease; ds.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9950454-A2.	
XX		
PD	07-OCT-1999.	
XX		
PF	26-MAR-1999; 99WO-US06473.	
XX		
PR	01-APR-1998; 98US-0054272.	
XX		
PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.	
XX		
PI	Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;	
XX		
DR	WPI; 1999-620066/53.	
DR	P-PSDB; AAY49570.	
XX		
PT	Determination of polymorphisms in genes, especially those identifying	
XX	predisposition to vascular disease	
XX	Claim 1; Fig 35; 134pp; English.	
CC	AAZ32159 to AAZ32194 represent reference alleles for specifically	
CC	claimed nucleic acid sequences from the present invention which comprise	
CC	polymorphic sites as given in a table in the specification, selected	
CC	from 92 single nucleotide polymorphisms in which the nucleotide at the	
CC	polymorphic site is different from a nucleotide at the same site in a	
CC	reference allele. The nucleic acids, and primers and probes, are used to	
CC	identify polymorphisms, which may predispose an individual to disease,	
CC	especially a vascular disease. They can also be used in phenotype	
CC	correlations, forensics, paternity testing, medicine or genetic	
CC	analysis. AAY49550 to AAY49573 represent the proteins which correspond	
XX	to some of the reference alleles.	
SQ	Sequence 3472 BP; 933 A; 817 C; 785 G; 937 T; 0 other;	
Alignment Scores:		
Pred. No.:	2.3e-37	Length: 3472
Score:	509.00	Matches: 120
Percent Similarity:	52.73%	Conservative: 83
Best Local Similarity:	33.17%	Mismatches: 140
Query Match:	21.83%	Indels: 42
DB:	20	Gaps: 14
US-09-208-629F-3 (1-407) x AAZ32191 (1-3472)		
Qy	20	LeuILEuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db	240	CTGCTGCTGGTGGCCGCCCTTC-----AGTCTGTGGCGCCGCTGTG 284
Qy	40	AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db	285	TCTGCCCGCACCCGGCGCCGAGCCAGCAATCAAAAGCAACAATGCCACCTTAGATCCC 344
Qy	53	LysSerPhe---AsnGlyGlyProGlnAenThrPheGluGluPheProLeuSerAspIle 71
Db	345	CGGTGCTATTTCTCTCAGGAACCCCAATGATAATATGAACCATTTGGGAGGATGAGGAG 404
Qy	72	GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 97
Db	405	AAAAATGAAGTGGGTAACTCAATACAGATTAGTCTCCATCAATAAAAGCAGTCCTCTT 464

Qy	88	AspSerIleSerThrLeuHisValAlaAsnAlaThrIleGlyTyrLeuArgSerSerLeu	107
Db	465	CAAAAACAACCTCTCTGCATTTCATCTCAGAAAGATGCCTCCGGATATTTTGACACAGCTCTCTGG	524
Qy	108	SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn	127
Db	525	CTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTTAAC	584
Qy	128	Ile-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIle	145
Db	585	ATCATGCCCATCGTTGTGTTCATCTCGAAATGAAGGTCAAGAAG---CCGGCGGTGGT	641
Qy	146	PheHisThrAsnLeuAlaIleAlaAapLeuPheCysValThrLeuProPheLysIle	165
Db	642	TACATGTGCACCTTGGCCAGCGAGATGTCTGTGTGTGTCTGCCCTTTAAGATC	701
Qy	166	AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr	185
Db	702	AGCTATTACTTTCCGGCAGTGATTTGGCATTTGGGTCTGAATTTGTCTCGTCTCACT	761
Qy	186	ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsn	205
Db	762	GCAGCAATTTTACTGTAACATAGTACGCTCTATCTTGTCTATGACAGTCATAAGCATTGAC	821
Qy	206	ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg	222
Db	822	CGSFTTCTGTGCTGTGGTATCCCATGACAGTCCCTCTCTCCGGGTACTCTCGGAAGGGCT	881
Qy	223	SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro	242
Db	882	TCTTTCACT-----TGTCTGGCCATCTGGCTTTGGCCATCGCAGGGGTAGTCCT	932
Qy	243	PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp	262
Db	933	CTCTGCTCTCAAGGAGCAACCATCCAGGTGCCGGCTCAACATCACTACCTGTCTCATGAT	992
Qy	263	ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla	282
Db	993	GTGCTCAATGAACCTGCTCGAAGGCTACTATGCCCTACTACTTCTCAGGCTTCTCTGCT	1052
Qy	283	PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis	302
Db	1053	GTCTTC--TTTITTTGGCCGCTCATCATTTCCACGGTCTGTTATGTGTCTATCATTCGA	1109
Qy	303	LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr	314
Db	1110	TGCTTACTGCTCTCCGCAAGTTGCCAACCGCAGCAAGAAGTCCGGGCTTTG-----TTC	1163
Qy	315	IleLysAlaValLeuLeuIleValIlePheThrIleCysPheAlaProThrAsnIle	334
Db	1164	CTGTCAGCTGCT-----GTTTCTGCATCTTCATCATTTTGTTCGAGCCCAACAAGCTC	1217
Qy	335	IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr	352
Db	1218	CTCTGATTGGCATTAC---TCATTCTTCTCACACTTCCACCACAGAGGCTGCCTAC	1274
Qy	353	PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu	372
Db	1275	TTTGCCTTACCTCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1334
Qy	373	TyrPheValMetSer	377
Db	1335	TACTATTACGCTTCC	1349
RESULT 23			
ID	AAZ50771		
XX	AAZ50771 standard; DNA; 3480 BP.		
AC	AAZ50771;		
XX	31-MAY-2000 (first entry)		
DT			
XX			

DE Human thrombin receptor DNA.

XX Human; thrombin receptor; Thr; PAR-1; protease activated receptor;

KW antisense molecule; PAR antibody; cytostatic; therapeutic;

KW metastatic tumour cell; placental implantation; invasive cell; ds.

XX Homo sapiens.

XX OS

XX Key Location/Qualifiers

PH 225..1502

FT /*tag= a

FT /product= "Thrombin receptor"

FT /transl_except= (pos:1500..1502, aa:Glx)

FT /note= "no stop codon given"

FT misc_binding

FT 320..570

FT /*tag= b

FT /bound_moiety= "Thr RNA probe"

XX WO200008150-A1.

XX 17-FEB-2000.

XX 05-FEB-1999; 99WO-IL00079.

XX 07-AUG-1998; 98IL-0125698.

XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

XX Bar-Shavit R;

XX WPI; 2000-205706/18.

XX P-PSDB; AAY45035.

XX Treating metastatic tumor cells useful for treating disorders involving

PT placenta implantation in a female comprises administration of an

PT antisense molecule complementary to an RNA sequence of a protease

PT activated receptor protein -

XX Example 2; Fig 1a; 46pp; English.

XX The patent discloses a method to treat metastatic tumour cells using

CC an antisense molecule comprising a polynucleotide complementary to an

CC RNA sequence of a protease activated receptor (PAR) protein, or an

CC antibody capable of binding to a PAR protein. The antisense molecules and

CC antibodies of PAR protein are also used to treat disorders associated

CC with implantation of placenta. The present sequence is a DNA encoding

CC human thrombin receptor (THR) (also known as PAR-1), which is a

CC G-coupled protein belonging to the PAR family. This sequence is used to

CC produce antisense molecules or antibodies against PAR protein for

CC therapeutic use.

XX SQ Sequence 3480 BP; 941 A; 817 C; 785 G; 937 T; 0 other;

Alignment Scores:

Pred. No.:	2,3e-37	Length:	3480
Score:	509.00	Matches:	120
Percent Similarity:	52.73%	Conservative:	83
Best Local Similarity:	31.17%	Mismatches:	140
Query Match:	23.83%	Indels:	42
DB:	21	Gaps:	14

US-09-208-629F-3 (1-407) x AAZ50771 (1-3480)

Qy 20 Leu1leuValAlaAlaGlyLeuLeuProValThrValCysGlnSerGlyIle 39

Db 240 CTGCTGTGTGGTGGCGGCTGCTTC-----AGTCTGTGGCGCGCTGTTG 284

Qy 40 AsnValSerAspAsnSerAlaIysPro-----ThrLeuThrIle 52

Db 285 TCTGCCCGCACCGGGCCCGGAGCCAGATCAAAAGCAACAAATGCCACCTTAGATCCC 344

Qy 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71

RESULT 24

AAK90981

ID AAX90981 standard; cDNA; 1534 BP.

Db 345 CGGTCACTTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAG 404

Qy 72 GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87

Db 405 AAAAATGAAGGGGTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGTCTCTTT 464

Qy 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTrpLeuArgSerSerLeu 107

Db 465 CAAAACAACACTTCTCGCATTCATCTCAGAAGATGCTCCGGATATTTGACCAGCTCTCGG 524

Qy 108 SerThrGlnValIleProAlaIleTrpIleLeuLeuPheValValGlyValProSerAsn 127

Db 525 CTGACACTCTTTGTCCCATCTGTACACCGGAGTGTGTGTAGTCAGCCTCCCACTAAAC 584

Qy 128 Ile-----ValThrLeuTrpLysSerLeuArgThrLysSerIleSerLeuValIle 145

Db 585 ATCATGGCCATCGTTGTGTTTCATCTCGAAATGAAGGTCAAGAAG---CCGGCGGTGGTG 641

Qy 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165

Db 642 TACATGCTGCACCTGGCCACGAGATGCTGTTGTGTCTGTCTGCCCTTTAAGATC 701

Qy 166 AlaTrpHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185

Db 702 AGCTATTACTTTCCGGCAGTGAATGGGTCTGAATTTGTGCTCGCTTCGCTCACT 761

Qy 186 ValValPheTrpGlyAsnMetTrpCysAlaIleLeuLeuThrCysMetGlyIleAsn 205

Db 762 GCAGCATTTTACTGTAAACATGACGCTCTATCTTCTCATCATGACATCAAGCATGAC 821

Qy 206 ArgTrpLeuAlaThrAlaHisPro-----PheThrTrpGlnLysLeuProLysArg 222

Db 822 CGGTTCTGGCTGGTGTATCCATGCAGTCCCTCTCTCGGGTACTCTGGGAAGGCT 881

Qy 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTrpMetLeuPro 242

Db 882 TCCTTCACT-----TGTCTGGCCATCTGGGCTTGGCCATCGCAGGGTAGTGCCT 932

Qy 243 PheValIleLeuLysGlnGluTrpHisLeuValHisSerGluIleThrCysHisAsp 262

Db 933 CTGCTCTCAAGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACCTGTCAATG 992

Qy 263 ValValAspAlaCysGluSerProSerPheArgPheTrpPheValSerLeuAla 282

Db 993 GTGCTCAATGAACCTGTCTCGAAGCTACTATGCTACTACTTCTCAGCCTTCTCTGCT 1052

Qy 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTrpThrThrLeuIleHis 302

Db 1053 GTCTTC---TTTTTTGTGGCGTGTATTCACCGTCTGTATGTGTATCATTCGA 1109

Qy 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTrp 314

Db 1110 TGTCTTAGCTTCTCCGAGTTGCCAACCGCAGCAAGAGTCCCGGCTTTG-----TTC 1163

Qy 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334

Db 1164 CTGTCACTGCT-----GTTTTCATCTTTCATCATTTGCTTCGACCCCAACAGTC 1217

Qy 335 IleLeuValIleHisAlaAsnTrpTrpHis-----AsnThrAspSerLeuTrp 352

Db 1218 CTCTGATTGCGCATTCAC---TCATTCTCTTCTCACACTTCCACACAGAGGTGCTCTAC 1274

Qy 353 PheMetTrpLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeu 372

Db 1275 TTTCCTACTCTCTGT 1334

Qy 373 TrpPheValMetSer 377

Db 1335 TACTATTACGCTCC 1349

14-MAR-2001 (first entry)
 Human low adenosine antisense oligonucleotide related sequence #2998.
 Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.
 Homo sapiens.
 WO200062736-A2.
 26-OCT-2000.
 24-MAR-2000; 2000WO-US08020.
 06-APR-1999; 99US-0127958.
 (UYEC-) UNIV EAST CAROLINA.
 (NYCE/) NYCE J W.
 Nyce JW;
 WPI; 2000-679539/66.
 Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -
 Disclosure; Page 1415; 1592pp; English.
 The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and chemokines, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.
 Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;
 Alignment Scores:
 Pred. No.: 1.09e-37 Length: 1534
 Score: 507.50 Matches: 111

Percent Similarity: 52.97% Conservative: 76
 Best Local Similarity: 31.44% Mismatches: 151
 Query Match: 23.76% Indels: 15
 DB: 21 Gaps: 6

US-09-208-629F-3 (1-407) x AAF21431 (1-1534)

QY 27 LeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAla 46
 DB 15 CTGCTCTCTGCGCCCTGTGTGTG-----GGTTGAGCTGTCTGGCGCACCCAG 65
 QY 47 LysProThrLeuThrIleLys-----SerPheAsnGlyGlyProGlnAsnThrPheGlu 64
 DB 66 ACCCCAGCGCTACGACGAGAGCGGAGCACCAGGAGTGGTGTATGATGACACGCGCTCA 125
 QY 65 GluPheProLeuSerAspIleGluGlyThrPheGlyAlaThrThrThrIleLysAlaGlu 84
 DB 126 ATCTGCTGCCCC-----CGCGCTACCCAGCCCAAGTC----- 161
 QY 85 CysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArg 104
 DB 162 TGTCCCAATGACAGT---GACACCTGGAGCTCCCGGACAGCTCAGGCGACCTGCTTCTG 218
 QY 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyVal 124
 DB 219 GGCTGGGTGCCACCAGGCTGTGCCCGCCCTCTATGGCTGTGTCTGTGTGGGGCTG 278
 QY 125 ProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
 DB 279 CCGGCCAATGGGCTGTGGGCTGTGGCCACGACGAGGACCTCGGCTGCCCTCCACC 338
 QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
 DB 339 ATGCTGCTGATGACCTCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 398
 QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 DB 399 ATCGCTTACACCTGCTGGCGCAGCGCTGCGGCGCGCGCTGCGGCGCGCTGCGCGCTG 458
 QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleThrCysMetGlyIle 204
 DB 459 ACGGCGCGACTCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 518
 QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 DB 519 GATCGTACTGCGCCCTGTGTGACCGCTGCGGCGCGCGCGCTGCGGCGCGCGCTG 578
 QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
 DB 579 GCCCTTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
 QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 DB 639 CTACAGCGGACAGCTTCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 698
 QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePhe 284
 DB 699 ---CCCTGGAGCGACAGGCTCCACCTGCGAACCGGCTTACCTGCTGCTGCTGCTGCTG 755
 QY 285 GlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeu 304
 DB 756 GGCTGTTCTGCGCCCTGTGTGGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 815
 QY 305 LysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIle 324
 DB 816 GCGGCGAGCGCGCGCTACGCGCCACGCGCTGAGCTGACCGAGTGGTGGTGGTGGTGGTGG 875
 QY 325 PheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyr 344
 DB 876 GCGGTGGCTTCTTGTGCGCCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 935
 QY 345 TyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeu 364

Db 793 GCCTTCCTGCTGCTC---TTTTTGTGCGCTGATCATTTCCACGGTCTGTATGTG 849
 Qy 299 ThrLeuIleHisIysLeuLysSer-----LysAspArgIle 310
 Db 850 TCTATCATTCGATGTCTTAGCTCTTCGCGAGTTGCCAACGCGACGAAGTCCCGGGT 909
 Qy 311 TrpLeuGlyTrpIleLysAlaValLeuLeuIleValIlePheThrIleCysPheAla 330
 Db 910 TTG-----TTCCTGTCAGTGTCT-----GTTTCTGCACTCTCATCTTCTCGGA 957
 Qy 331 ProThrAsnIleLeuValIleHisHisAlaAsnTrpTrpTrpHis-----AsnThr 348
 Db 958 CCCACAAAGCTCTCTGATTCGCAATTAC--TCATCTCTTCTCACACTTCCACACACA 1014
 Qy 349 AspSerLeuTrpPheMetTrpLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
 Db 1015 GAGCTGCTCTACTTTGCT 1074
 Qy 369 AspProPheLeuTrpPheValMetSer 377
 Db 1075 GACCCCTAATTACTATTAGCTTCC 1101

RESULT 31

AAF21432
 ID AAF21432 standard; DNA; 3182 BP.

AC AAF21432;
 XX
 XX
 DT 14-MAR-2001 (first entry)
 XX
 XX
 DE Human low adenine antisense oligonucleotide related sequence #2999.

XX Low adenine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammation;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; ROS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.

XX Homo sapiens.

XX WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenine (A) content antisense oligonucleotides which do not
 PT trigger adenine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -

XX Disclosure; Page 1415-1416; 1592pp; English.

XX The present invention describes low adenine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the

CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adenosine molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Alignment Scores:
 Pred. No.: 9,21e-37 Length: 3182
 Score: 502.00 Matches: 113
 Percent Similarity: 54.15% Conservative: 76
 Best Local Similarity: 32.38% Mismatches: 130
 Query Match: 23.50% Indels: 30
 DB: 21 Gaps: 12

US-09-208-629F-3 (1-407) x AAF21432 (1-3182)

Qy 49 ThrLeuThrIleLysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPhePro 67
 Db 43 ACCTTAGATCCCGGTCACTTTCTCTCAGAACCCCAATGATAAATATGAACCATTTGG 102
 Qy 68 LeuSerAspIleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAla 83
 Db 103 GAGGATGAGGAGAGAAATAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAA 162
 Qy 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTrpLeu 103
 Db 163 AGCAGTCTCTTCAAAAACAACTTCTCGCATCTCAGAGATCCCTCCGGAATTTTG 222
 Qy 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyriIleLeuPheValValGly 123
 Db 223 ACCAGTCTCTGGCTGACACTTTTGTCCCATCTGTGTACACCGGAGTGTGTAGTCAGC 282
 Qy 124 ValProSerAsnIle-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
 Db 283 CTCCCACTAAACATCATGCGCATCGTTGTGTTCATCTCGAAATGAAGTCAAGAAG--- 339
 Qy 142 SerLeuValIlePheHisThrAsnLeuAlaIleAspLeuLeuPheCysValThrLeu 161
 Db 340 CCGCGGTGTGTGTACATGTGCACCTGGCCACGCGAGATGTGTGTGTGTGTGTGTCTC 399
 Qy 162 ProPheLysIleAlaTyriHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys 181
 Db 400 CCCTTTAAGATCAGTATTACTTTTCCGCGCAGTGTGGCAGTTGGTCTGAATTTGTGT 459
 Qy 182 ArgIleThrThrValValPheTyriGlyAsnMetTyriCysAlaIleLeuLeuLeuThrCys 201
 Db 460 CGTTCGTCTCAGTCAGCATTTTACTGTAAACATGTACGCTCTATCTTGTCTCATGACAGC 519
 Qy 202 MetGlyIleAsnArgTrpLeuAlaThrAlaHisPro-----PheThrTrpGlnLys 218
 Db 520 ATAAGCATTTGACCGGTTTCTGGCTGTGGTGTATATCCATGCAGTCCCTCTCTGGGTACT 579
 Qy 219 LeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeu 238

Db 580 CTGGGAAGGGCTCTCTTCACT-----TGCTGCCCATCTGGGCTTTGGCCATGCGCA 630
 Qy 239 TyrMetLeuProPheValIleLeuLeuGlnGluTyrHisLeuValHisSerGluIleThr 258
 Db 631 GGGGTAGTGCCTCTCGTCTCAAGGAGCAACCATCCAGTGCCTGGGCTCAACATCACT 690
 Qy 259 ThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPhe 278
 Db 691 ACCTGTATGATGTGCTCAATGAACCTGCTCGAAGCTACTATGCTACTACTACTCTCA 750
 Qy 279 ValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleIleIlePheCysTyrThr 298
 Db 751 GCCTTCTCTGCTCTTC-----TTTTTGTGCCGTGATCAATTCACGGTCTGTATGTG 807
 Qy 299 ThrLeuIleHisLeuLeuSer-----LysAspArgIle 310
 Db 808 TCTATCATTCAGTCTTACTCTTCCTCGCAGTTGCCAACCGCAGCAAGAGTCCCGGCT 867
 Qy 311 TrpLeuGlyTyrIleLeuAlaValLeuLeuLeuValIlePheThrIleCysPheAla 330
 Db 868 TTG-----TTCCTGTGAGTGT-----GTTTCTGCATCTTCATCACTTTGCTTGGGA 915
 Qy 331 ProThrAsnIleIleValIleHisHisAlaAsnTyrTyrThrHis-----AsnThr 348
 Db 916 CCACAAACGTCCTCTGATGGCAATAC---TCATTCCTTTCTCACTTCCACCCACA 972
 Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
 Db 973 GAGGCTGCCTACTTTGGCTACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
 Qy 369 AspProPheLeuTyrPheValMetSer 377
 Db 1033 GACCCCTAATTTACTATTACGCTTC 1059

RESULT 32

AA35310
 ID AA35310 standard; DNA; 3182 BP.

XX
 AC AA35310;

XX 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:184.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US17712.

XX 03-AUG-1998; 98US-0095212.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or

PT cancers -

XX Disclosure; Page 1330-1331; 1343pp; English.

XX The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1880
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.

XX Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Alignment Scores:

Pred. No.: 9,21e-37 Length: 3182
 Score: 502.00 Matches: 113
 Percent Similarity: 54.15% Conservative: 76
 Best Local Similarity: 32.38% Mismatches: 130
 Query Match: 23.50% Indels: 30
 DB: 21 Gaps: 12

US-09-208-629F-3 (1-407) x AAA35310 (1-3182)

Qy 49 ThrLeuThrIleLysSerPhe---AenGlyGlyProGlnAsnThrPheGluGluPhePro 67
 Db 43 ACCTTAGATCCCGGTCAITTTCTCTCAGAACCCCAATGATAAATATGAACCAATTTGG 102
 Qy 68 LeuSerAspIleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAla 83
 Db 103 GAGGATGAGGAG 162
 Qy 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
 Db 163 AGCAGTCTCTTCAAAAACAACCTTCTGTCATCTCAGAGATCGCTCCGATATTTC 222
 Qy 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
 Db 223 ACCAGCTCTGGCTGACACTTTTGTCCCATCTGTGTACACCGAGTGTGTGTGTGTGTGT 282
 Qy 124 ValProSerAsnIle-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIle 141
 Db 283 CTCCCAATAACATCATGCGCATCTGTTGTCTCATCTCTGAAATGAGGTCAAGAG--- 339
 Qy 142 SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
 Db 340 CCGGCGGTGTGTACATGCTGACCTGGCCACGCGCAGATGTGTGTGTGTGTGTGTGTGT 399
 Qy 162 ProPheLysIleAlaTyrHisLeuAenGlyAsnAsnTrpValPheGlyGluValMetCys 181
 Db 400 CCCTTTAAGATCAGCTATTACTTTTCCCGCAGTGATTGGCAGTGTGGGTCTGAATGTGT 459
 Qy 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCys 201
 Db 460 CGCTTCGTCACCTGCAGCATTTTACTGTAAATGTACGCTCTATCTTGTCTCATGACGATC 519

Db 576 AACGTGGTACCGTGGCGCTTTTACGCAACATGATTCCAGCATCTCCACCATCACTGCT 635
 Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
 Db 636 ATCAGCGTGGAGCGCTTCTGGGGTCTGTACCGCTCAGTCCAAGCGTGGCGCGC 695
 Qy 222 ArgSerPheSerLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeu 241
 Db 696 CGTGTATACGGGTGGCGCGTGTGTGAGGACCTGGCTGTCTCTCAGCGCGCTTCC 755
 Qy 242 ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHis 261
 Db 756 CCGTGGCGCCACCGATCTCACCTACCGGTGACCGCTGGGCATCATCACCCTGCTC 815
 Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
 Db 816 GACGTCTCAAGTGCAGCATCTCCCGCAGCGTGGCCATGTGGCGGTGCTCTCTTCCACC 875
 Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
 Db 876 ATCTTCATCTCTGTG-----TTCTCATCCCGTTTGTGTGATCACCAGTGGCTGT 923
 Qy 297 TyrThrThrLeuIleHisLysLeuLysSerLysAspArgIleTrp----- 311
 Db 924 TACAGCGCCACCATCTCAAGCTTTGCGACGGAGGAGCGCACCGCGGGAGCAGCGG 983
 Qy 312 LeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaPro 331
 Db 984 AGCGCGCGGTGGCGCTGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1043
 Qy 332 ThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeu 351
 Db 1044 AACAACTTCGTCTCTGGCGCACATCGTGAGCGCGCTGTCTTACGGC---AAGAGCTAC 1100
 Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 371
 Db 1101 TACCAGGTGACAAGCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1160
 Qy 372 LeuTyrPheValMetSerLys 378
 Db 1161 GTTATTACTTTGGTCCCGG 1181
 RESULT 36
 AAD01131
 ID AAD01131 standard; cDNA; 1080 BP.
 XX AC AAD01131;
 XX DT 02-NOV-2000 (first entry)
 XX DE Human orphan G protein-coupled receptor hCHN4 cDNA.
 XX KW Human; orphan G protein-coupled receptor; GPCR; hCHN4; drug screening;
 XX KW transmembrane receptor; expressed sequence tag; EST; signal cascade; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..1080
 FT /*tag= a
 FT /product= "hCHN4"
 FT /note= "Human orphan G protein-coupled receptor"
 XX FN WO200031258-A2.
 XX PD 02-JUN-2000.
 XX PF 13-OCT-1999; 99WO-US23687.
 XX PR 20-NOV-1998; 98US-0109213.
 XX PR 16-FEB-1999; 99US-0120416.
 XX PR 26-FEB-1999; 99US-0121852.
 XX PR 12-MAR-1999; 99US-0123946.

PR 12-MAR-1999; 99US-0123949.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0136567.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 28-JUN-1999; 99US-0141448.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156634.
 PR 23-SEP-1999; 99US-0156853.
 PR 01-OCT-1999; 99US-0157280.
 PR 01-OCT-1999; 99US-0157281.
 PR 01-OCT-1999; 99US-0157282.
 PR 01-OCT-1999; 99US-0157293.
 PR 01-OCT-1999; 99US-0157294.
 PR 12-OCT-1999; 99US-0416760.
 PR 12-OCT-1999; 99US-0417044.
 XX XX
 PA (AREN-) ARENA PHARM INC.
 XX PI Chen R, Dang HT, Liaw CW, Lin I;
 XX DR WPI; 2000-400068/34.
 DR P-PSDB; AAY71304.
 XX PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
 XX PS for use in the identification of G protein-coupled receptor agonists -
 XX PS Claim 53; Page 76; 102pp; English.
 XX CC The present sequence is a cDNA encoding hCHN4, an endogenous human
 CC orphan G protein-coupled receptor (GPCR). The hCHN4 cDNA was identified
 CC using full length EST (expressed sequence tag) 1184934 as a probe.
 CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
 CC alpha helices with an extracellular N-terminus and an intracellular
 CC C-terminus. However, no endogenous ligands has yet been identified for
 CC the proteins of the invention. The orphan GPCRs may be used in the
 CC identification of their endogenous ligands, and to screen potential GPCR
 CC agonists and antagonists for use as pharmaceutical agents. The proteins
 CC may also be used in the study of GPCR-mediated signalling cascades, and
 CC to elucidate their precise role in normal and diseased human conditions.
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue
 CC localisation expression analysis to provide information about their
 CC function in healthy and pathological states.
 XX SQ Sequence 1080 BP; 162 A; 369 C; 322 G; 227 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,25e-36 Length: 1080
 Score: 494.00 Matches: 106
 Percent Similarity: 53.54% Conservative: 53
 Best Local Similarity: 35.69% Mismatches: 116
 Query Match: 23.13% Indels: 22
 DB: 21 Gaps: 6
 US-09-208-629F-3 (1-407) x AAD01131 (1-1080)
 Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115
 Db 28 GACAACGCGACGCTGCAGATGCTGCGGAACCCCGCGATCGCGTGGCCCTGCGCGTGTG 87
 Qy 116 TyrIleLeuLeuPheValValProSerAsnIleValThrLeuTyrLysLeuSer 135
 Db 88 TACTCGTGTGGCGCGGTGACATCCCGGGCAACCTCTCTCTGTGGGTGTGTCG 147
 Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154
 Db 148 CGCGCGATGGGCGCCAGATCCCGTCCGTCACTTTCATGATCACTGACCTGACCGGAC 207
 Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 174

Db 208 CTGATGCTGGCCAGCGTGTGGCTTTCCAAATCTACTACCATGCAACGCCACCACTGG 267
Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194
Db 268 GTATTCGGGTGCTGCTTTGCAACGCTGGTACCGTGGCCCTTTACGCAACATGTTATCC 327
Qy 195 AlaIleLeuIleLeuThrCysMetGlyLeuAsnArgTyrLeuAlaThrAlaHisProPhe 214
Db 328 AGCATCTCACCATGACCTGATCAGCGTGGAGCGCTCTCTGGGGTCTGTACCGCTC 387
Qy 215 ThrTyrGlnIlyLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGTCCACGCGTGGCGCGCGCTGTTACCGCGTGGCGCGTGGCGGACCTGGCTG 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnIleTyrHisLeuValHis 254
Db 448 CTGCTCTGACCGCCCTGTGGCCCTGGCGGCGCACCGATCTACCTACCGGTGACGCC 507
Qy 255 SerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSer----- 271
Db 508 CTGGGCATCATCCTGCTTCGAGCTCTCAAGTGGAGCATGCTCCCGCGGTGGCCATG 567
Qy 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
Db 568 TGGGCGGTGTTCTCTTCACCATCTTCATCTGCTG-----TTCTCATCCCG 615
Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeu----- 304
Db 616 TTCTGTATCAGCGGTGTTTACCGCCACCATCTCAAGCTGTGGCGAGGAG 675
Qy 305 -----LysSerLysAspArgIleTyrPheGlyTyrIleLysAlaValLeuLeuIle 321
Db 676 GCGCACGCGCGGAGCAGCGGCGGTGGCGCTGGCGGTGGTCTTGTGCTG--- 732
Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAla 341
Db 733 -----GCCTTTGTACCTGCTTCGCCCCAACCACTTCGTGCTCTCTGGCGCACATCGT 786
Qy 342 AsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeu 361
Db 787 AGCGCGCTGTCTACGGC---AAGAGCTACTACACGCTGTACAGCTCAGCTGTGTCTC 843
Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 844 AGCTGCTCAACAACCTGCTGGACCGTGTGTTTATTACTTTGCGTCCCG 894

RESULT 37
AAA46032
ID AAA46032 standard; cDNA; 1080 BP.
AC AAA46032;
XX
XX 22-AUG-2000 (first entry)
XX
XX Human G protein coupled receptor hCHN4 encoding cDNA SEQ ID NO:29.
XX
XX Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW identification; agonist; screening; therapeutic; pharmaceutical;
KW mutant; SS.
XX
XX Homo sapiens.
XX
XX WO200022131-A2.
XX
XX 20-APR-2000.
XX
XX 13-OCT-1999; 99WO-US24065.
XX
XX 13-OCT-1998; 98US-0170496.
PR 12-NOV-1998; 98US-0108029.
PR 20-NOV-1998; 98US-0109213.
PR 27-NOV-1998; 98US-0110060.
PR 16-FEB-1999; 99US-0120416.

PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123944.
PR 12-MAR-1999; 99US-0123945.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123948.
PR 12-MAR-1999; 99US-0123949.
PR 12-MAR-1999; 99US-0123951.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 28-MAY-1999; 99US-0137567.
PR 30-JUN-1999; 99US-0141448.
PR 27-AUG-1999; 99US-0151114.
PR 03-SEP-1999; 99US-0152524.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156634.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX
XX WPI; 2000-317986/27.
DR P-PSDB; AAB02838.
XX
XX Non-endogenous, human G protein-coupled receptors for screening
PT receptor, inverse or partial agonists useful as therapeutic agents -
XX
XX Example 1; Page 105-106; 187pp; English.
XX
XX The present invention describes transmembrane receptors, preferably
CC human G protein coupled receptors (GPCR), for which the endogenous
CC ligand is unknown (orphan GPCR receptors). More specifically the present
CC invention relates to non-endogenous, constitutively activated versions
CC of a human GPCR. These non-endogenous human GPCRs can be useful for
CC the direct identification of candidate compounds as receptors agonists,
CC inverse agonists or partial agonists for use as pharmaceutical agents.
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC the exemplification of the present invention.
XX
XX Sequence 1080 BP; 162 A; 369 C; 322 G; 227 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1,25e-36 Length: 1080
Score: 494.00 Matches: 106
Percent Similarity: 53.54% Conservative: 53
Best Local Similarity: 35.69% Mismatches: 116
Query Match: 23.13% Indels: 22
DB: 21 Gaps: 6
US-09-208-629f-3 (1-407) x AAA46032 (1-1080)
Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115
Db 28 GACAAGCGACGCTGCAGATGCTGCGGAACCCCGCGATCGCGTGGCCGCGTGGTG 87
Qy 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeuSer 135
Db 88 TACTCGCTGCTGGCGGCTCAGCATCCCGGGCAACCTCTCTCTCTGCTGCTGCTGTC 147
Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154
Db 148 CGGCGCATGGGCGCCAGATCCCGCTGGTCACTTCATGATCAACCTGACCGCGACG 207
Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyr 174
Db 208 CTGATGCTGGCCAGCGTGTGCTTTCCAAATCTACTACCATGCAACGCCACCACTGG 267
Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194

Db 268 GTATTGGGGTGGCTTTCGAACGTGGTGACCGTGGCCCTTTTACGCCAAACATGTATTCC 327

Qy 195 AlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214

Db 328 AGCATCTCACCAGTATGATCAGCGTGGAGCGCTTCTGGGGGTCTGTACCGCTC 387

Qy 215 ThrTyrGlnIysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpVal 234

Db 388 AGCTCCAAAGCGTGGCCCGCGTGTACGCGTGGCCGCGTGTGAGGAGCCTGTGCTG 447

Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHis 254

Db 448 CTGCTCTCAGCGCCCTGTGCGCCGCGGCGACCGCATCTCCTACCGGTGACGCC 507

Qy 255 SerGluLeuThrCysHisAspValValAspAlaCysGluSerProSer----- 271

Db 508 CTGGGCATCATCCTGCTTCGACGTCTCAAGTGGACGATGCTCCCGCGGTGGCCATG 567

Qy 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289

Db 568 TGGCGCGTGTCTCTTCCACCATCTTCATCTGCTG-----TTCTCATCCCG 615

Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeu----- 304

Db 616 TTCGTGATCAGCGTGGCTGTTCACGCGCCACCATCTCAAGCTGTTCGCGAGGAG 675

Qy 305 -----LysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuIle 321

Db 676 GCGCACGCGGGAGCAGCGAGGCGCGGTGGCGCTGGCGCGGTGCTTCTGCTG--- 732

Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAla 341

Db 733 -----GCCTTTGTACCTGCTTCGCGCCCAACACTTCTGCTCTCGCGGCACATCGTG 786

Qy 342 AsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeu 361

Db 787 AGCCGCTGTCTACGCG-----AAGAGCTACTACACGCTGACAGCTCAGCTGTGTCTC 843

Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378

Db 844 AGCTGCTCAACAACCTGCTGGACCGCTTGTATTATTACTTTGGTCCCGG 894

RESULT 38

AAAX90979

ID AAAX90979 standard; cDNA; 1361 BP.

XX AC AAAX90979;

XX DT 25-JAN-2000 (first entry)

XX DE Mouse protease-activated receptor 4 (PAR4) cDNA.

XX KW Mouse protease-activated receptor 4 cDNA; PAR4; thrombin receptor; phosphoinositide hydrolysis; calcium efflux; platelet aggregation; platelet activation; thrombin agonist; thrombin antagonist; therapeutic; wound; blood coagulation; heart attack; stroke; inflammatory response; proliferative response; atherosclerosis; restenosis; glomerulosclerosis; pulmonary inflammation; ARDS; adult respiratory distress syndrome; ds.

XX OS Mus musculus.

XX FH Key

XX FT 1..1191

XX FT CDS

XX FT /*tag= a

XX FT /product= "PAR4"

XX W09943809-A2.

XX PD 02-SEP-1999.

XX PF 11-FEB-1999; 99WO-US02983.

XX PR 27-FEB-1998; 98US-0032397.

XX (REGC) UNIV CALIFORNIA.

XX Coughlin SR, Kahn M;

XX WPI; 1999-619953/53.

XX P-PSDB; AAY15081.

XX DNA molecules encoding protease-activated receptor 4, useful in compound assays for thrombin agonist and antagonist activity -

XX Claim 3; Fig 1; 69pp; English.

XX The present sequence is a cDNA encoding protease-activated receptor 4 (PAR4) and derived from a mouse 14-15 days embryo library. PAR4 is a G protein-coupled thrombin receptor expressed on cell surface. It is activated by thrombin and mediates signalling events e.g. phosphoinositide hydrolysis, calcium efflux and platelet aggregation. The receptor is highly expressed in spleen cells and likely to be involved in thrombin-mediated activation of platelets and other haematopoietic cells. It is used for screening novel thrombin agonists or antagonists. The agonists are used as therapeutics to treat wounds, promote clotting and as reagents to activate platelets in diagnostic tests. Antagonists are used to control blood coagulation, treat heart attacks and strokes, and block inflammatory and proliferative responses that occur in normal wound healing and variety of diseases including atherosclerosis, restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.

XX Sequence 1361 BP; 234 A; 416 C; 394 G; 317 T; 0 other;

Alignment Scores:

Pred. No.:	1.69e-36	Length:	1361
Score:	494.00	Matches:	105
Percent Similarity:	53.03%	Conservative:	70
Best Local Similarity:	31.82%	Mismatches:	139
Query Match:	23.13%	Indels:	16
DB:	20	Gaps:	5

US-09-208-629F-3 (1-407) x AAAX90979 (1-1361)

Qy 48 ProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPhePro 67

Db 127 CCCACAGTAGAATCTCAAGAG-----CCGAGTCTCTCAGACAGCTTAATCCA 174

Qy 68 LeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGlu 87

Db 175 -----CGAGGCTACCCGCGCAAAATTC-----TGTGCCAAC 204

Qy 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107

Db 205 GACAGT--GACAGCTGGAGCTCCCGCGCAGCTCTCAAGCAGCTGTCTGGGGTGGTA 261

Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127

Db 262 TCCACAGGCTGATGACTCTGCTCTATGGCTTGTGGTGGCTGTGGGCTGCTGCAAT 321

Qy 128 IleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHis 147

Db 322 GGGCTGGCGCTGGGTGCTGGCCACAGGGTGGCCAGCTGCTCCCATCTTGCTC 381

Qy 148 ThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyr 167

Db 382 ACGAACCTGGCAGTGGCTGATTCGCTGTTGGCCCTGTGGCCGCCACACAGCTGGCTTAC 441

Qy 168 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValVal 187

Db 442 CACTTGGTGGCGCAGCGCTGGCCATTTGGTAGAGTGTGCTGGGGTGGCCACAGTGCC 501

Qy 188 PheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyr 207

Db 502 CTCATATGGCCACATGATGTTAGTGTGCTGCTGGTGCAGTGCAGTGGACAGATAC 561

QY 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
 Db TGGGCCGTGTCTCTTCACCATCTTCATCTGCTG-----TTCTCATCCCG 615
 QY 290 PheVallellellePheCysTyrThrThrLeuLeuHissylsLeu----- 304
 Db TTCTGATACCGTGGCTTTTACACGCCACCATCTCAAGCTGTGGCCACGGAGGAG 675
 QY 305 -----LysSerTyrAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 321
 Db GCGCACGCCGGAGGAGCGCGCGGTGGCGCTGCGCGGTGGTCTTGTCTG--- 732
 QY 322 LeuVallellellePheCysPheAlaProThrAsnIlelleleuVallelleHissAla 341
 Db -----GCCTTGTGTCACCTCTTGGCCCCCAACAACCTCTGCTCTCTGGGGCACGTGTG 786
 QY 342 AsnTyrTyrTrpHisAsnThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeu 361
 Db AGCGCGCTGTCTACGGC---AGAGCTACTACACGCTGACAGCTCAAGCTGTGTCTC 843
 QY 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db AGCTGCTCAACAACCTCTGCGACCGCTTGTGTTTATTACTTTGGTCCCG 894

RESULT 40
 AAI64231
 ID AAI64231 standard; cDNA; 1080 BP.
 XX
 AC AAI64231;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE G-protein coupled receptor (designated Paul) encoding sequence.
 XX
 KW G-protein coupled receptor; paul; 7TM receptor; GPC receptor;
 KW human; antibacterial; viricide; fungicide; protozoicide;
 KW anti-Human Immunodeficiency Virus; analgesic; cytosolic; antidiabetic;
 KW anorectic; osteopathic; antiparkinsonian; cardiant; hypertensive;
 KW hypotensive; osteopathic; antianigmal; antimanic; cerebroprotective;
 KW antitumor; antiallergic; antidepressant; antimigraine; antiemetic;
 KW tranquiliser; neuroleptic; neuroprotective; nootropic; anticonvulsant;
 KW Human Immunodeficiency Virus type 1; HIV-1; HIV-2; pain; cancer;
 KW diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease;
 KW acute heart failure; hypertension; hypertensive; urinary retention;
 KW osteoporosis; angina pectoris; myocardial infarction; stroke; ulcer;
 KW allergy; benign prostatic hypertrophy; migraine; vomiting; anxiety;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1080
 FT /*tag= a
 FT /product= "Paul protein"
 FT /note= "G-protein coupled receptor referred to as Paul
 FT in the specification"
 XX
 XX
 PN US2001029032-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 18-DEC-2000; 2000US-0739151.
 XX
 PF 09-OCT-1998; 98US-103789P.
 PR 06-OCT-1999; 99US-0413534.
 XX
 XX (ZHUY/) ZHU Y.
 PA (LIXX/) LI X.
 PA (VAWT/) VAWTER L.
 XX
 XX Zhu Y, Li X, Vawter L;
 FI
 XX

DR WPI; 2001-647985/74.
 DR P-PSDB; AAG78530.
 XX
 PT New G-protein coupled receptor polypeptide, referred as Paul and
 PT encoding polynucleotide, useful for diagnosing and treating cancers,
 PT infections, neurological disorders, diabetes, asthma and identifying
 PT modulators
 XX
 PS Claim 2b; Page 12-13; 15pp; English.
 XX
 CC The invention relates to an isolated polypeptide, a member of G-protein
 CC coupled receptor family of polypeptides, comprising a fully defined
 CC sequence of 359 amino acids, its 95% identical sequence, or a polypeptide
 CC encoded by a polynucleotide comprising a fully defined sequence of 1080
 CC base pairs defined in the specification, or a fragment or variant of it.
 CC The activity of the protein of the invention may be described as
 CC antibacterial, viricide, fungicide, protozoicide, anti-Human
 CC Immunodeficiency Virus, analgesic, cytosolic, antidiabetic, anorectic,
 CC osteopathic, antiparkinsonian, cardiant, hypertensive, hypotensive,
 CC osteopathic, antianigmal, antimanic, cerebroprotective, antitumor,
 CC antiallergic, antidepressant, antimigraine, antiemetic, tranquiliser,
 CC neuroleptic, neuroprotective, nootropic and anticonvulsant. Polypeptides
 CC and polynucleotides of the invention are useful in diagnosis and in
 CC identifying compounds such as agonists and antagonists which are useful
 CC in therapy. They are also useful for treating diseases, including
 CC infections such as bacterial, fungal, protozoan and viral infections,
 CC particularly infections caused by Human Immunodeficiency virus type 1 or
 CC 2 (HIV-1 or HIV-2), pain, cancers, diabetes, obesity, anorexia, bulimia,
 CC asthma, Parkinson's disease, acute heart failure, hypotension,
 CC hypertension, urinary retention, osteoporosis, angina pectoris,
 CC myocardial infarction, stroke, ulcers, allergies, benign prostatic
 CC hypertrophy, migraine, vomiting, psychotic and neurological disorders,
 CC including anxiety, schizophrenia, manic depression, depression, delirium,
 CC dementia, and severe mental retardation, dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome. Polypeptides and
 CC polynucleotides of the invention may also be useful as diagnostic
 CC reagents, for example in the detection of mutations in the associated
 CC gene, chromosome localisation studies and expression pattern
 CC determination. The current sequence represents a G-protein coupled
 CC receptor (designated "Paul" in the specification) encoding sequence.
 XX
 SQ Sequence 1080 BP; 162 A; 370 C; 321 G; 227 T; 0 other;

Alignment Scores:
 Pred. No.: 1,54e-36 Length: 1080
 Score: 493.00 Matches: 106
 Percent Similarity: 53.54% Conservative: 53
 Best Local Similarity: 35.69% Mismatches: 116
 Query Match: 23.08% Indels: 22
 DB: 23 Gaps: 6

US-09-208-629F-3 (1-407) x AAI64231 (1-1080)

QY 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrClnValleProAlaIle 115
 Db GACACCGCAGCTGCAGATGCTGCGAACCAGCGGATCGCGTGGCCCTCCCGGTG 87
 QY 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSer 135
 Db TACTCGCTGGTGGCGGCGGTGACATCCCGGGCAACCTCTTCTCTGTGGGTGCTGTC 147
 QY 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAsp 154
 Db CGCGCGATGGGGCCAGATCCCGTGGTCACTTCATGATCACTGACCTGACGTCACGGAC 207
 QY 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrp 174
 Db CTGATGCTGGCCAGCGTGTGCTTTTCAAAATCTACTACCACTTACTACCAACCCACTGG 267
 QY 175 ValPheGlyGluValMetCysArgIleThrValValPheThrGlyAsnMetTyrCys 194
 Db GTATTCCGGGTGCTGCTTTTGGCAACGTGGTGGCGCTTTTACGCCAAACATGATATCC 327

```
Qy 195 AlaIleLeuIleThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214
Db 328 AGCATCTCTACCATGACTGATACAGCGTGGAGCGCTTCTGGGGTCTCTGTACCCGCTC 387
Qy 215 ThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGCTCCAAGCGCTGGCGCGCTGCTTACGGCGTGGCGCGTGTGCAGGGAGCTGGCTG 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHis 254
Db 448 CTGCTCTGACCGCCCTGTCTCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCC 507
Qy 255 SerGluIleThrThrCysHisAspValAlaAspAlaCysGluSerProSer----- 271
Db 508 CTGGGCATCATCCTCTTCGACGCTCTCAAGTGGAGCATGTCTCCCGAGCGTGGCCATG 567
Qy 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
Db 568 TGGGCGGTGTCTCTTTCACCATCTTCATCTCTGCTG-----TTCTCATCCCG 615
Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeu----- 304
Db 616 TTCGTGATACCGGTGGCTGTGTACCGGCCACCATCTCAAGCTGTTCGCGACGGAGGAG 675
Qy 305 -----LysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuIle 321
Db 676 GCGCACGGCGGGAGCGAGCGGCGCGGTGGGCTGTGGCGGTGTCTGTCTG--- 732
Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAla 341
Db 733 -----GCCTTTGTACCTGCTTCGCCCCCAACAACTTCGTGCTCTGTGGCGCACATCGTG 786
Qy 342 AsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeu 361
Db 787 AGCCGCTGTCTCTACGGC---AAGAGCTACTTACCACGCTGTACAGCTCACGCTGTGTCTC 843
Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 844 AGCTGCCCTCAACAACACTGTCTGGACCCGTTGTATTATTACTTTGGTCCCGG 894
```

Search completed: June 29, 2003, 08:18:33
Job time : 243.264 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:29:54 ; Search time 2264.22 Seconds
(without alignments)
5231.316 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTYXQHPVAGSQDKMKIL.....AMARPLXRRPRDIWIHAW 407

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101625_12922/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CGN 1 1 2827 @runat_24062003_101625_12922 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2121	99.3	1224	6	AR070424 Sequence
2	2121	99.3	2409	10	MMU92972
C 3	1929	90.3	235097	2	AC110877 Mus muscu
4	1729.5	81.0	1505	10	AF310076 Rattus no
5	1443.5	67.6	1224	6	AR070426 Sequence
6	1443.5	67.6	1830	9	HSU92971 Human prote
7	1372.5	64.3	9273	9	AF374726 Homo sapi
C 8	1372.5	64.3	101887	9	AC026725 Homo sapi
9	1372.5	64.3	131370	9	AC026706 Homo sapi
10	1369	64.1	1102	6	AR070427 Sequence
11	1300.5	60.9	1124	6	AR070425 Sequence
12	584	27.3	2713	10	MMPAR2MR
13	584	27.3	2732	6	AR012639 Sequence
14	584	27.3	2732	6	AR0171259 Sequence
15	584	27.3	2732	6	187850 Sequence 60
16	583	27.3	2772	10	BC025432 Mus muscu
17	582.5	27.3	1680	5	XLU09632 Xenopus lae
18	581.5	27.2	1477	10	MMPAR2CR2
C 19	581.5	27.2	199227	2	AC110876 Mus muscu
20	579.5	27.1	1475	6	AR012637 Sequence
21	579.5	27.1	1475	6	AR171257 Sequence
22	579.5	27.1	1475	6	142454 Sequence 1
23	579.5	27.1	1475	6	187848 Sequence 1
24	569	26.6	2813	9	BC012453 Homo sapi
25	568	26.6	2876	9	BC018130 Homo sapi
C 26	564.5	26.4	1428	10	RNU61373 Rattus norv
27	564.5	26.4	151166	2	AC130634 Rattus no
C 28	564	26.4	52368	9	AC010621 Homo sapi
29	563	26.4	1255	6	AR012638 Sequence
30	563	26.4	1255	6	AR171258 Sequence
31	563	26.4	1255	6	142455 Sequence 3
32	563	26.4	1255	6	187849 Sequence 3
33	563	26.4	1289	9	HSPAR2B Human prote
34	563	26.4	1451	9	HSU34038 Homo sapi
35	563	26.4	18351	9	AF400075 Homo sapi
C 36	563	26.4	184536	2	AC068682 Homo sapi
37	563	26.4	204485	2	AC114962 Homo sapi
38	560.5	26.2	1124	9	HSU36753 Human prote
C 39	559.5	26.2	88955	2	AC128225 Rattus no
40	551	25.8	1414	6	AR012640 Sequence
41	551	25.8	1414	6	AR171260 Sequence
42	551	25.8	1414	6	187851 Sequence 62
43	540.5	25.3	3313	10	BC031516 Mus muscu
44	534.5	25.0	199227	2	AC110876 Mus muscu
45	533	25.0	5620	10	MMTHRECO2 Mus musculu

ALIGNMENTS

RESULT 1

AR070424
LOCUS AR070424 1224 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5892014.
ACCESSION AR070424
VERSION AR070424.1 GI:7221312
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1224)
Coughlin,S.R., Ishihara,H. and Connolly,A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 1.06-APR-1999;
FEATURES Location/Qualifiers
source 1. 1224
/organism="unknown"
BASE COUNT 299 a 341 c 239 g 345 t
ORIGIN
Alignment Scores:
Pred. No.: 2.6e-163 Length: 1224
Score: 2121.00 Matches: 404
Percent Similarity: 99.26% Conservative: 0
Best Local Similarity: 99.26% Mismatches: 3
Query Match: 99.30% Indels: 0
DB: 6 Gaps: 0
US-09-208-629F-3 (1-407) x AR070424 (1-1224)
Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
Db 3 ACTTTGATATCTTAACAACATCCCTGTAGCCGGGTCTCAGGACATCAAGATGAAATCCTT 62
Qy 21 IleLeuValAlaAlaGlyLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
Db 63 ATCTTGGTGTGAGCTGGCTGCTGTTCTGCCAGTCACCTGTTTCCCAAGTGGCATAAAT 122
Qy 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
Db 123 GTTTCAGACCACTCAGCAACCAACCTTAACCTAATTAAGAGTTTATGGGGTCCCAA 182
Qy 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyThrGlyAlaThrThr 80
Db 183 AATACCTTTGAAGAATTTCCCACTTCTGCATAGAGGCTGGACAGAGGCCACCAACT 242
Qy 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100
Db 243 ATAAAGCGGAGTGTCCCGAGGAGCAGTATTTCAACTCTCCACGGAATAATGTACCATA 302
Qy 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
Db 303 GGATACCTGAGAAGTCTCTTAAGTACCAAGTATACCTGCCATCTATATCTGCTGTT 362
Qy 121 ValValGlyValProSerAsnIleValThrLeuThrLysLeuSerLeuArgThrLysSer 140
Db 363 GTGGTGTGTGTACCATCAACATCGTACCCCTGTGGAACCTCTCTTAAGGACCAAAATCC 422
Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 423 ATCAGTCTGGTCATCTTTTCCACCAACCTGGCCATCGCAGATCTCTTTCTGTGTGACA 482
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 483 CTGCCATTTAAGATCGCTACCATCTCAATGGCAACCACTGGGTATTTGGCGAGGTATG 542
Qy 181 CysArgIleThrThrValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
Db 543 TGCCTGGATCACCGGTCGTTTCTACGCCAACATGTACTCGCTCATCTCTCATCCTCACT 602
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 603 TGCAATGGGCATCAACCGCTACTGGCCACGGCTCACCTTTCATACCAAGAGCTGCC 662

Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 663 AAACGCAGCTTCTCTTCT 722
Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
Db 723 CTGCCCTTTGTCTATCTCTGAAGCAGGAGTACCACTCGTCCACTCAGAGATCACCACTGC 782
Qy 261 HisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
Db 783 CACGATGTCTGACCGCGTGGAGTCCCATCATCTCTCGGATCTCTACTACTCTCTCTCT 842
Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
Db 843 TTAGCATTTCTTTGGGTTCTCTCATCCGTTTGTGATCATCATCTCTCTGTACAGACTCTC 902
Qy 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
Db 903 ATCCACAAATTTAAATCAAAAGGATCGGATATGGCTGGCTACATCAAGGCGCTCTCTCT 962
Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
Db 963 ATCTTTGTGATTTTCAAAATTTGCTTTGCCCCCAACATCATCTCTGTAATCCACCAT 1022
Qy 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCys 360
Db 1023 GCCAACTACTACTTACCACAAATACCGACAGCTTGATCTTTATGTATCTTATTGCTCTG 1082
Qy 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
Db 1083 CTGGGAGGCTGAATAGCTGCCTAGATCCATTCCTTTACTTTGTCTCATGTGCAAGTGTGA 1142
Qy 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
Db 1143 GATCAGCTTAATCTTCTAGTCGCAATGGCAAGCACTTTAGAGACCAAGGAGATATC 1202
Qy 401 TrpGluAspIleHisAlaTrp 407
Db 1203 TGGGAAGACATACATCTCTGG 1223
RESULT 2
MMU92972 2409 bp mRNA linear ROD 16-APR-1997
LOCUS Mus musculus protease-activated receptor 3 (PAR3) mRNA, complete
DEFINITION cds.
ACCESSION U92972
VERSION U92972.1 GI:1938376
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2409)
AUTHORS Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
Timmons,C., Tram,T. and Coughlin,S.R.
TITLE Protease-activated receptor 3 is a second thrombin receptor in
humans
JOURNAL Nature 386 (6624), 502-506 (1997)
MEDLINE 97242411
PUBMED 9087410
REFERENCE 2 (bases 1 to 2409)
AUTHORS Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
Timmons,C., Tram,T. and Coughlin,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-1997) CVRI, UCSF, 3rd and Parnassus, San
Francisco, CA 94143, USA
FEATURES Location/Qualifiers
source 1. .2409
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
1. .2409
/gene="PAR3"

CDS 153..1262
 /gene="PAR3"
 /note="thrombin receptor; coagulation protease"
 /codon_start=1
 /product="protease-activated receptor 3"
 /protein_id="AAC53137.1"
 /db_xref="GI:1938377"
 /translation="MKILILVAAGLLFLPVTVCOSGINVSDNSAKPTLTITKSNFGPQ
 LTFEEPLSDIEGWTGATTTIKAECPEDSISTLHVNNATIGYLRSLSTQVIPIAYIL
 LPVGVPSNVTIWKLSRTKSLIVIFHTNLAIADLLFCVTLPFKIAYHLNNGNVF
 GYVCRITTVVFGNNMYCAILITICGYNRYLATAHPFTYOKLPKPSFSLGCGIYVW
 MVLYMLPPVYLQEVHLVHSEITTCDDVVDACESPSSFEFYFVSLAFEGFLIPVI
 IICVYTLHLKLSKRILWGYKAVULLIVIFTICFAPNTIILVIHHANYHYHTDS
 LYFMYLIALCGSLNSCLDPFLYFVMSKVVDQLNP"
 BASE COUNT 598 a 652 c 454 g 703 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 5,55e-163 Length: 2409
 Score: 2121.00 Matches: 404
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 10 Gaps: 0

US-09-208-629F-3 (1-407) x MMU92972 (1-2409)

QY 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
 DB 105 ACTTTGTATTAACAACATCTCTGTAGCGGGTCTCAGACATCAAGATGAATACTCTT 164
 QY 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
 DB 165 ATCTTGGTTCAGCTGGGCTGCTTTCTGCCAGTCACTGTTGCCAAAGTGGCATAAAT 224
 QY 41 ValSerAspSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
 DB 225 GTTTCAGAACACTCAGCAAGACCACTTAACATTAAGAGTTTTAATGGGGGTGCCAA 284
 QY 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80
 DB 285 AATACCTTTGAGAGATTTCCACTTTCTGACATAGAGGGCTGGACAGAGGCCACCACT 344
 QY 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100
 DB 345 ATAAAGCGGAGTGTCCCGAGGACAGATTTCAACTCTCCACGTGAATAATGCTACCATA 404
 QY 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
 DB 405 GGATACCTGAGAAGTTCCTTAAGTACCAAGTACCAAGTACCTGCTATATCTCTGCTGTT 464
 QY 121 ValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer 140
 DB 465 GTGGTTGGTGATACCATCAACATCGTGACCTGTGGAACTCTCTTTAAGAGCAAAATCC 524
 QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 DB 525 ATCAGTCTGTGTCATCTTTTACACCAACTGGCCATCGCAGATCTCTCTTTCTGTGTACA 584
 QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 DB 585 CTGCAATTTAAGATCGGCTACCATCTCAATGGCAACAACTGGGTATTTTGGCGAGGTGATG 644
 QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
 DB 645 TGCCGGATCACCGGTCGTTTCTACGGCAACATGACTGCGCTATCTGTATCTCTCACT 704
 QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 DB 705 TGCATGGGCATCAACCGCTACCTGGCCACGGCTACCTCTTCATACACAGAGGTGCC 764
 QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240

DB 765 AAACGCAGCTTCTCCTTGCTCATGTGTGGCATAGTGTGGTCAATGGTTTCTTATACATG 824
 QY 241 LeuProPheValIleLeuLysGlnGlyTyrHisLeuValHisSerGluIleThrThrCys 260
 DB 825 CTGCCCTTTGTCTATCTGAAGCAGGAGTACCACCTCTGTCACCTCAGAGATCACCACCTGC 884
 QY 261 HisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
 DB 885 CAGATGTCTGCACGCGTGCAGAGTCCCATCATCTCTCCGATTTCTACTACTTCTGCTCC 944
 QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 DB 945 TTAGCATTTCTTGGGTCTCTCATCCCTTTGTGATCATCATCTTCTGTACAGACTCTC 1004
 QY 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
 DB 1005 ATCCACAACCTTAATCAAGGATCGATATGCTGGCTACATCAAGGCCGCTCTCTC 1064
 QY 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
 DB 1065 ATCCTTGTGATTTTCAAAATTTGCTTTGCCCCACCAACATCATCTCGTAAATCCACCAT 1124
 QY 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleIleAlaLeuCys 360
 DB 1125 GCCAACTACTACTACCACAATACCGACAGCTTGTACTTATGTATCTTATGTCTGTGTC 1184
 QY 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
 DB 1185 CTGGGAGGCTGATAGTGCCTAGATCCATCTCTTACTTGTCTGTCGAAAGTTGTA 1244
 QY 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
 DB 1245 GATCAGCTTAATCTTCTAGTCGCAATGGCAAGACCACTTTAGAGACCAAGAGAGATATC 1304
 QY 401 TrpGluAspIleHisAlaTrp 407
 DB 1305 TGGGAAGACATACATCTGCTGG 1325

RESULT 3
 AC110877 235097 bp DNA linear HTG 16-FEB-2002
 AC110877/c
 LOCUS Mus musculus clone RP23-345121, WORKING DRAFT SEQUENCE, 3 unordered
 DEFINITION pieces.
 AC110877
 AC110877.1 GI:18693419
 VERSION HTG: HTGS PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 235097)
 Martin, J., Hosseini, R., Peng, Y., Peng, Z., Rubin, E.M. and Cheng, J.F.
 Direct Submission
 Unpublished
 2 (bases 1 to 235097)
 Martin, J., Hosseini, R., Peng, Y., Peng, Z., Rubin, E.M. and
 Cheng, J.-F.
 Direct Submission
 Submitted (16-FEB-2002) Berkeley PCA, Lawrence Berkeley National
 Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
 America
 Draft Sequence Produced by Berkeley PGA
 pga.lbl.gov
 Center Code: PGABERK.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2480: contig of 2480 bp in length

Query Match: 80.97% Indels: 6
DB: 10 Gaps: 3
US-09-208-629F-3 (1-407) x AF310076 (1-1505)

Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
Db 21 TCTCTGACACTTAGTGACATCCCATAAACCGGGTCTCAAGACATGAGATGAAGTCTTT 80
Qy 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle--- 39
Db 81 ATCTTGGTGGGGTAGGCTGCTGTTCTGCCGACCACTGTTTCCCAAAGTGGCATGAAA 140
Qy 40 AsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyPro 59
Db 141 CATGTTTCAGACAACTCAGCC-----TTAACTGCTGAGAGCTTTTAATGGCAAC-- 188
Qy 60 GlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 79
Db 189 GAACATTCTCTTGAAGAAATTCCTTGAAGTACCAAGATGATCTGCCATCTACATCTCGT 248
Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThr 99
Db 249 ACTATAAAGCGAGTGTCCGAGAAAGCAATTACAACCTCTCCATGTGATATATGCTACC 308
Qy 100 IleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeu 119
Db 309 ATGGGATACCTGAGAAGTTCCTTAAGTACCAAGATGATCTGCCATCTACATCTCGTGTG 368
Qy 120 PheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys 139
Db 369 TTTGTGATGGTGTACCAGCGAACACTCGTGACCCGTGGAACTCTCTCTCAAGGACCAAA 428
Qy 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
Db 429 TCCATCTGTCTGGTCATCTTTTACCAACCACTGGCCATCGCGATCTCTTTCTGTGTC 488
Qy 160 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
Db 489 ACGCTGCGGTTTAAGATCGCTACCATCTCAATGGGAACGACTGGGTCTTTGGCGAGTC 548
Qy 180 MetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeu 199
Db 549 ATGTGCGGGGTACACCGGTGCTTCTACGGCAACATGATTTGTGTATTTCTCATCTC 608
Qy 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
Db 609 ACCTGTATGGCATCAACCGCTACTGCGCACGGTCCATCTCTTCATCATACCGCAAGCTG 668
Qy 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
Db 669 CCCAAACGCAACTTCATTTGCTCATGTGTGGCGTGGTGGTTCATGGTGTCTTATAT 728
Qy 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThr 259
Db 729 ATGTGCGGTTGGCCATCTCTCAAGCAGGAGTACCATCTTGTCCAAACAGGGATCACCACC 788
Qy 260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheVal 279
Db 789 TGGCACAGCTGCCAGACACATGGAGTCCCGTGGCTTCCAGTCTTACTACTCTGCTC 848
Qy 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThr 299
Db 849 TCCTTGGCTCTTTGGGTCTTATCCCATTTGGTGTGCTGCTGCTTCTGTACACGACT 908
Qy 300 LeuIleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeu 319
Db 909 CTCATCACAAAGCTTAACCGCAGGATCGCAAAATGGGTGAGGTACATCAAGGCGGTTCTC 968
Qy 320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis 339
Db 969 CTCATCTTGTGATTTTTCACCATCTGCTTTGGCCCCCAACCAATCATCATATAATTCAC 1028

340 HisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeu 359
1029 CACGCCAACTACTACTACAGCAACACTGATAGCTTGTACTTATGTATCTCATAGCTCTC 1088
Qy 360 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysVal 379
1089 TGCTTGGGAGCGCTGAATAGTTCCTAGACCCCATCTCTTTATTTATCATGTCAAAAT 1148
Qy 380 ValAspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAsp 399
1149 GTAGATCAGCTGACCTCGTAGTCAACAATGGCGAGATCATCTTGGAGACAAGATAGAT 1208
Qy 400 IleTrp-GluAspIleHisAlaTrp 407
1209 ATCTGGGGAATGTATATGCTTGG 1233

RESULT 5
AR070426
LOCUS AR070426 1224 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 4 from patent US 5892014.
ACCESSION AR070426
VERSION AR070426.1 GI:7221314
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE 1 (bases 1 to 1224)
AUTHORS Coughlin, S.R., Ishihara, H. and Connolly, A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 4 06-APR-1999;
FEATURES
Location/Qualifiers
1..1224
source
BASE COUNT 303 a 316 c 221 g 384 t
ORIGIN

Alignment Scores:
Pred. No.: 2,54e-108 Length: 1224
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 6 Gaps: 1

US-09-208-629F-3 (1-407) x AR070426 (1-1224)

Qy 11 GlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeuLeuPheLeu 30
Db 40 GGGACTCAGGTCTCAATAAATGAAAGCCCTCATCTTTGCAGCTGCTGGCTCTCTGCTCTG 99
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 100 TTGCCACACTTTTGTCTGAGTGGCATGGAAATGATACAAACAACTTGGCAAGCCCAACC 159
Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 160 TTACCATTAAGACCTTTCGTGGAGCTCCCAAAATCTTTTGAAGAGTTCCTCTTCT 219
Qy 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 220 GCCTTGGAAAGGCTGGACAGGACGACGATTTACTGTAAAAAATTAAGTCCCTCGAAGAA 279
Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 280 GCTTTCATCTCCATGTGAAAAATGTACCATGGGTACCTGACAGCTCTTAAAGTACT 339
Qy 110 GlnValIleProAlaIleTyrIleLeuPheValIleGlyValProSerAsnIleVal 129
Db 340 AACTGATACCTGCCATCTACCTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 399
Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 400 ACCCTGTGGATGCTTTTCTTTCAGGACGACATCCATCTGTACCATCTGTATTTCTACACCA 459


```

Qy      210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
Db      727 ATCTCCATCTTTTACACTACCGGGCTGCCCCAAGCACACCTATGCTTGGTAACATGT 786
Qy      230 GlyTleValTrpValMetValPheLeuTyrMetLeuProPheValLleLeuLysGlnGlu 249
Db      787 GGACTGTGTGGGCAACAGTTTCTTATATATGCTGCCATTTTTCATCACTGAAGCAGGAA 846
Qy      250 TyrHisLeuValHisSerGluLeuThrThrCysHisAspValValAspAlaCysGluSer 269
Db      847 TATATCTTGTTCAGCCAGACATCACCACTGCCATGATGTTCAACACACTTGGAGTCC 906
Qy      270 ProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
Db      907 TCATCTCCCTCCCAACTTATCTACTTCTCTGCTGTCATCTTTTGGATCTCTTAATTCCA 966
Qy      290 PheValLleLleLlePheCysTyrThrThrLeuLleHisLysLeuLysSerLysAspArg 309
Db      967 TTTGTGCTTATCATCTACTGCTATGCGAGCATCATCCGACACTTAATGATACATCAT 1026
Qy      310 IleTrpLeuGlyTyrLleLysAlaValLeuLeuLleLeuValLlePheThrLleCysPhe 329
Db      1027 AGATGGTGTGGTATGTATAGCGAGCTCTCTCATCTCTTGATTTTACATTTGCTTT 1086
Qy      330 AlaProThrAsnLleLleLeuValLleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
Db      1087 GCTCCAAGCAATATTATTCTTATTATTCACCATGCTCACTACTACTACAACACACTGAT 1146
Qy      350 SerLeuTyrPheMetTyrLeuLleAlaLeuLysLeuGlySerLeuAsnSerCysLeuAsp 369
Db      1147 GGCTTATATTTTATATATCATCATAGCTTTGTGCTGGGTAGTCTTAATAGTTGCTTAGAT 1206
Qy      370 ProPheLeuTyrPheValMetSerLys 378
Db      1207 CCATTCCTTTATTTTCTCATGTCAAAA 1233

RESULT 7
AF374726      9273 bp  DNA  linear  PRI 01-JUN-2001
LOCUS      Homo sapiens coagulation factor II receptor-like 2 (F2RL2) gene,
DEFINITION      complete cds.
ACCESSION      AF374726
VERSION      AF374726.2 GI:14278710
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 9273)
AUTHORS      Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
TITLE      Direct Submission
JOURNAL      Submitted (28-APR-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REFERENCE      2 (bases 1 to 9273)
AUTHORS      Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
TITLE      Direct Submission
JOURNAL      Submitted (01-JUN-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REMARK      Sequence update by submitter
COMMENT      On Jun 1, 2001 this sequence version replaced gi:13991934.
            To cite this work please use: SeattleSNPs, NHLBI Program for
            Genomic Applications, UW-PHCRC, Seattle, WA (URL:
            http://pga.mbt.washington.edu).
FEATURES
source      1..9273
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
repeat_region      120..221
            /rpt_family="L1MB3"
            /rpt_type="dispersed"

```

```

variation      419
            /frequency="0.29"
            /replace="G"
variation      483
            /frequency="0.02"
            /replace="T"
variation      506
            /frequency="0.32"
            /replace="G"
variation      507
            /frequency="0.25"
            /replace="C"
variation      512
            /frequency="0.24"
            /replace="A"
variation      535
            /frequency="0.24"
            /replace="T"
variation      937
            /frequency="0.02"
            /replace="C"
gene      969..7322
            /gene="F2RL2"
mRNA      join(969..1176,5701..7322)
            /gene="F2RL2"
            /product="coagulation factor II receptor-like 2"
variation      1007
            /gene="F2RL2"
            /frequency="0.02"
            /replace="G"
            join(1113..1176,5701..6761)
            /gene="F2RL2"
            /note="thrombin receptor-like 2"
            /codon_start=1
            /product="coagulation factor II receptor-like 2"
            /protein_id="AAK51564.1"
            /db_xref="GI:13991935"
            /translation="MKALIFAAAGLLLLPTFCQSGMENDTNNLAKPTLPKTFRGAP
            PMSFEFFPSALEGWTGATITVKIKCPESASHLVKNATMGVLTSSLSATKLPALVYL
            LVFVGVPANAVTLMVLFRTSICCTVETNLAIADELFCVTLPKIAYHLNGNMV
            FEVLCRATVIFYGNMYSILLACISINRYLAIVHPFTYRGLPKHTVALVTCGLVM
            ATVLYMLPFFILKOEYLVQPDITTCSDVHNTCESSSPQLYYFISLAFPGELIPFV
            LIICYAAIIRITNAYDHRWLWYKASLLILVITFCFAPSNIILIIHHANYNNNTD
            GLYFIYLIALCLGSLNSCLDPFLYFLMSKTRHSTAYLTK"
variation      1156
            /gene="F2RL2"
            /frequency="0.01"
            /replace="C"
variation      1430
            /gene="F2RL2"
            /frequency="0.06"
            /replace="G"
variation      2055
            /gene="F2RL2"
            /frequency="0.03"
            /replace="T"
variation      2120
            /gene="F2RL2"
            /frequency="0.01"
            /replace="T"
variation      2136
            /gene="F2RL2"
            /frequency="0.01"
            /replace="G"
variation      2168
            /gene="F2RL2"
            /frequency="0.02"
            /replace="T"
variation      2174
            /gene="F2RL2"
            /frequency="0.01"
            /replace="-"
repeat_region      2250..2441

```


Db 6276 TGTGGACTGGTGGGCAACAGTTTCTATATATGCTGCCATTTTTCATCTAAGCAG 6335
Qy 249 GluTyrHisLeuValHisSerGluLeuThrThrCysHisAspValValAspAlaCysGlu 268
Db 6336 GAATATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTTCAACACACTTGGAG 6395
Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeu 288
Db 6396 TCCTCATCTCCCTTCCAACTCTATTACTTCTATCTCTTGGCATTTTGGATTCTTAAT 6455
Qy 289 ProPheValLeuLeuLeuPheCysTyrThrThrLeuLeuHisLysSerLysAsp 308
Db 6456 CCATTGTGCTTATCATCTACTGCTATGCAGCCATCATCCGACACTTAATGATAGAT 6515
Qy 309 ArgLeuTrpLeuGlyTyrLeuLeuAlaValLeuLeuLeuValLeuPheThrLeuCys 328
Db 6516 CATAGATGGTGTGTGATGTTAAGCGAGTCTCTCATCTTGTGATTTTACCATTTGC 6575
Qy 329 PheAlaProThrAsnLeuLeuLeuValLeuHisLysHisAlaSerTyrTyrHisAsnThr 348
Db 6576 TTTGCTCCAAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACACT 6635
Qy 349 AspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 6636 GATGGCTTATTTTATATATCTCATAGCTTTGCTGGTGTCTTAATAGTTGCTTA 6695
Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 6696 GATCATCTCTTATTTTTCATGTCAAA 6725

RESULT 8
AC026725/c
LOCUS 101887 bp DNA linear PRI 21-JUL-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2236F14, complete sequence.
ACCESSION AC026725
VERSION AC026725.7 GI:14993696
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 101887)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 101887)
TITLE DOE Joint Genome Institute.
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 101887)
TITLE Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 4 (bases 1 to 101887)
TITLE Submitted (27-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 5 (bases 1 to 101887)
TITLE Submitted (21-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
COMMENT On Jul 21, 2001 this sequence version replaced gi:14572130.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.1.
STS Content:
SHGC-102125 G56565
SHGC-101620 G56147
SHGC-83998 G53567.
FEATURES
Location/Qualifiers
1. .101887
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2236F14"
BASE COUNT 29200 a 20058 c 21105 g 31524 t
ORIGIN
Alignment Scores: 2.09e-100 Length: 101887
Pred. No.: 1372.50 Matches: 248
Score: 83.71% Conservative: 45
Percent Similarity: 70.86% Mismatches: 54
Best Local Similarity: 64.26% Indels: 3
Query Match: 9 Gaps: 1
DB:
US-09-208-629F-3 (1-407) x AC026725 (1-101887)
Qy 29 PheLeuProValThrValCysGlnSerGlyLeuAsnValSerAspAsnSerAlaLysPro 48
Db 93492 TTTCTTTCAATTACAGGCATGGAA-----AATGATACAAACAACACTTGGCAAGCCA 93442
Qy 49 ThrLeuThrLeuLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 93441 ACCTTACCATTAAGACCTTTCTGTGGAGCTCCCCAAATTCCTTTGAAGAGTTTCCCTTT 93382
Qy 69 SerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAlaGluCysProGluAsp 88
Db 93381 TCTGCTTGGAGGCTGGACAGGACGACGATTACTGTAAATAATTAAAGTCCCTGAAGAA 93322
Qy 89 SerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSer 108
Db 93321 AGTGCTTCACATCTCCATGTGAAAAATGTACTACATGGGTACCTGACCACTCTTAAAGT 93262
Qy 109 ThrGlnValIleProAlaIleTyrIleLeuPheValGlyValProSerAsnIle 128
Db 93261 ACTAAACTGATACCTGCCATCTACCTCTGGTGTGTGTAGTTGGTCTCCCGGCAATGCT 93202
Qy 129 ValThrLeuTrpLysSerLeuArgThrLysSerIleSerLeuValIlePheHisThr 148
Db 93201 GTGACCTGTGGATGCTTTTCTTCAGACACAGATCCATCTGTACCATCTGATTCTACACC 93142
Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHis 168
Db 93141 AACCTGGCATTGCAGATTTCTTTTGTGTGTACATTCCTTTTAAAGTAGCTTATCAT 93082
Qy 169 LeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPhe 188
Db 93081 CTCAATGGGAACAACCTGGGTATTGTGAGAGGTCTGTGCGGCGCCACACAGTCATCTTC 93022
Qy 189 TyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeu 208
Db 93021 TATGGCAACATGATCTGCTCCATCTGCTCTGCTGCTGATCAGCATCAACCTGCTACCTG 92962
Qy 209 AlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMet 228
Db 92961 GCCATCGTCCATCTTTCACCTACCGGGCTGCCCAAGCACACCTATGCTTGGTAACA 92902
Qy 229 CysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGln 248
Db 92901 TGTGGACTGGTGGGCAACAGTTTCTTATATATATGCTGCCATTTTTCATCTGAAAGCAG 92842
Qy 249 GluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAlaCysGlu 268
Db 92841 GAATATTATCTTGTTCAGCCAGACATCACCCTGCTGATGATGTTTCAACACACTTGGAG 92782
Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeu 288
Db 92781 TCCTCATCTCCCTTCCAACTCTATTACTTCTATCTCTGTCATCTTCTTGGATTCTTAAT 92722
Qy 289 ProPheValIleIleIlePheCysTyrThrThrLeuLeuHisLysLysSerLysAsp 308
Db 92721 CCATTTGTGCTTATCATCTACTGCTATGACGCCATCATCCGGACACTTAATGATCAGCAT 92662

```

Qy 309 ArgileTrpLeuGlyTyrIleLeuValLeuLeuLeuValIlePheThrIleCys 328
Db 92661 CATAGATGGTGTGGTATGTTAAGCGAGCTCTCTCATCTTGTGATTTTACCAATTGC 92602

Qy 329 PheAlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrHisAsnThr 348
Db 92601 TTTGCTCCAGCAATATTATTCTTATTATTACCATGCTAACTACTACTACAACACT 92542

Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 92541 GATGGCTTATATTTTATATATCTCATAGCTTTGTGCTGGTAGTCTTAATAGTTGCTTA 92482

Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 92481 GATCCATTCCTTTATTTTCTCATGTCAAAA 92452

RESULT 9
AC026706 131370 bp DNA linear PRI 21-JUN-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2068N7, complete sequence.
DEFINITION AC026706
ACCESSION AC026706
VERSION AC026706.5 GI:14518404
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 131370)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
REFERENCE 2 (bases 1 to 131370)
DOE Joint Genome Institute.
Direct Submission
Unpublished
REFERENCE 3 (bases 1 to 131370)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 131370)
DOE Joint Genome Institute and Stanford Human Genome Center.
Submitted (21-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jun 21, 2001 this sequence version replaced gi:13677033.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.3.
STS Content:
WI-11362 G24506.
FEATURES
Location/Qualifiers
1..131370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2068N7"
BASE COUNT 40462 a 26452 c 25592 g 38854 t
ORIGIN
Alignment Scores:
Pred. No.: 2,78e-100 Length: 131370
Score: 1372.50 Matches: 248
Percent Similarity: 83.71% Conservative: 45
Best Local Similarity: 70.86% Mismatches: 54
Query Match: 64.26% Indels: 3
DB: 9 Gaps: 1
US-09-208-629F-3 (1-407) x AC026706 (1-131370)

Qy 29 PheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAlaLysPro 48
Db 60186 TTTCTTTCAATTACAGCATGAA-----AATGATACAAACAACTTGGCAAGCCA 60236

```

```

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 60237 ACCTTACCATTAAAGACCTTTTGGTGGAGCTCCCCCAATCTTTTGAAGAGTTCCCTTT 60296

Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 60297 TCTGCCTTGAAGGCTGGACAGGAGCCACGATTACTGTAAAAATTAAGTGCCTGAAGAA 60356

Qy 89 SerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 108
Db 60357 AGTGCCTTCACATCTCCATGTGAAATGTACCAGGGTACCTGACCAAGCTCTTAAAGT 60416

Qy 109 ThrGlnValIleProAlaIleTyrIleLeuPheValValGlyValProSerAsnIle 128
Db 60417 ACTAAACTGATACCTGCCATCTACCTCTGGTGTGTGTAGTTGGTGTCCCGGCCAATGCT 60476

Qy 129 ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr 148
Db 60477 GTGACCTGTGGATGCTTTTCTTCAGGACCAATCCATCTGTACCACCTGTATTCTACACC 60536

Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHis 168
Db 60537 AACCTGGCCATTGCAGATTCTTTTGTGTGTACATTTGCCCTTTAAGATAGATTATCAT 60596

Qy 169 LeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPhe 188
Db 60597 CTCATTTGGGACAACTTGGGTATTTGGAGAGGCTCTGTCCGGGCCACCAAGTCACTTC 60656

Qy 189 TyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeu 208
Db 60657 TATGGCAACATGATCTCTCCATTCTGCTTCTTGTGCTGTCATCAGCATCAACCGCTACCTG 60716

Qy 209 AlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuMet 228
Db 60717 GCATTCGTTCATCTTTCACCTACCGGGGCTGCCCAAGCACCACTATGCTTGGTAACA 60776

Qy 229 CysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGln 248
Db 60777 TGTGGACTGGTGGGCAACAGTTTCTTATATATGCTGCCATTTTTCATCTGAAGCAG 60836

Qy 249 GluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGlu 268
Db 60837 GAATATTATCTTGTTCAGCCAGACATCACCACTGCCATGATTTTCAACAACACTTGCAG 60896

Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPheGlyPheLeuIle 288
Db 60897 TCTCATCTCCCTCCAACTTATTACTTCTCTCTGTCATCTCTTGGATCTTTGATTTTAAT 60956

Qy 289 ProPheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAsp 308
Db 60957 CCATTTGTGCTTATCATCTACTGCTATGCGAGCCATCATCGGACACTTAATGATCAGCAT 61016

Qy 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
Db 61017 CATAGATGGTGTGGTATGTTAAGGCGAGTCTCTCTCATCTCTGTGATTTTACCATTTGC 61076

Qy 329 PheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrHisAsnThr 348
Db 61077 TTTGCTCCAAAGCAATATTATTCTTATTATTCACATGCTAACTACTACTACAACACT 61136

Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 61137 GATGGCTTATATTTTATATATCTCATAGCTTTTGTGCTGGTAGTCTTAATAGTTGCTTA 61196

Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 61197 GATCCATTCCTTTATTTTCTCATGTCAAAA 61226

RESULT 10
AR070427 AR070427 1102 bp DNA linear PAT 18-FEB-2000
LOCUS DEFINITION Sequence 5 from patent US 5892014.
ACCESSION AR070427

```

```
VERSION AR070427.1 GI:7221315
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1102)
AUTHORS Coughlin,S.R., Ishihara,H. and Connolly,A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 5 06-APR-1999;
FEATURES
    source
        1..1102
            /organism="unknown"
BASE COUNT 276 a 286 c 195 g 345 t
ORIGIN
Alignment Scores:
Pred. No.: 2,52e-102 Length: 1102
Score: 1369.00 Matches: 245
Percent Similarity: 84.96% Conservative: 43
Best Local Similarity: 72.27% Mismatches: 51
Query Match: 64.09% Indels: 0
DB: 6 Gaps: 0
US-09-208-629F-3 (1-407) x AR070427 (1-1102)
Qy 40 AsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyPro 59
Db 13 AATGATACAAACAACTTGGCAAGCCAACTTACCCATTAAAGACCTTTCGTGGAGCTCC 72
Qy 60 GluAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 79
Db 73 CCAAAATCTTTTGAAGAGTTCCCTTTTCCTTGGAGGCTGGACAGGACCGATT 132
Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThr 99
Db 133 ACTGTAAATAATTAGTCCCTCGAAGAAAGTCTTCACATCTCCATGGAATAATGCTACC 192
Qy 100 IleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeu 119
Db 193 ATGGGGTACCTGACCAGCTCTTAAGTACTAACTGATGATACCTGCCATCTACCTCGTGTG 252
Qy 120 PheValValGlyValProSerAsnIleValThrLeuThrIleLysSerLeuArgThrLys 139
Db 253 TTTGTAGTTGGTGTCCGGCAATGCTGTGACCTGTGGATGCTTTTCTTCAGGACCCAGA 312
Qy 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
Db 313 TCCATCTGTACCCTGATTTCTACACCACTGGCCATTGCAGATTTCTTTTGTGTT 372
Qy 160 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
Db 373 ACATTGGCCCTTAAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTATTTGGAGAGTC 432
Qy 180 MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeu 199
Db 433 CTGTGCGGGCCACCAAGTCATCTTCTATGGCAACATGACTGCTCCATCTGCTCCTT 492
Qy 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
Db 493 GCCTGCATCAGCATCAACCGCTACTGGCCATGCTCCCTTTTCACCTACCGGGCGCTG 552
Qy 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
Db 553 CCCAAGCACACCTATGCTTGGTAACATGTGGACTGTGGTGGCAACAGTTTCTTATAT 612
Qy 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThr 259
Db 613 ATGCTGCATTTTTCATACCTGAAGCAGGAATATTATCTTGTTCAGCCAGACATCACCACC 672
Qy 260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheVal 279
Db 673 TGCCATGATGTTCAACAACCTTGGAGTCTCTCATCTCCCTTCCAACTCTATTACTTCATC
```

```
Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHi 168
Db 299 AACCTGCCCATCGGGATCTCTTTCTGTGTACAGTGTGGCTTAAAGATCNCCCTACCA 358
Qy 168 sLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValph 188
Db 359 TCTCAATGGCAACAACCTGGGTATTTGGCGAGTTCATGTGGCGGATCACACGGTGCCTTT 418
Qy 188 eTyrGlyAsnMetTyrCysAla----IleLeuIleLeuThrCysMetGlyIleAsnArgTy 207
Db 419 CTACGGCAACATGATCTACGCGTANNNTCCGTATCCCTCCTGATGGCGATCAACCGCTA 478
Qy 207 rLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLe 227
Db 479 CCTGGCACGGCTCACCTTTCACATACAGAGCTGCCAAGCGAGCTTCTCCATGCT 538
Qy 227 uMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuL 247
Db 539 CATGTGTGGCATGTGTGGGTCTATGTTTCTTATACATGCTGCCCTTTGTATCCNNNA 598
Qy 247 ysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaC 267
Db 599 AGCAGGAGTACACCTCGTCCACTCCGAGATCACACCTGCGCAGATGTCTGTCACCGCT 658
Qy 267 ysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheL 287
Db 659 CGGANTCCCATCATCTCTCCGATCTACTACTTCTCTCTCTAGCATTTCTTGGGTTC 718
Qy 287 euIleProPheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerL 307
Db 719 TCATCCCGTTGTGATCATCATCTCTCTGTACACGACTCTCATCCACAACTTAAATCAA 778
Qy 307 ysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrI 327
Db 779 AAGATCGATATGCTGGCTGATCAAGCGCGTCTCTCTCTCTGATTTGTAATTTACCA 838
Qy 327 leCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHisA 347
Db 839 TCTGCTTCCCCCACCACAG----- 858
Qy 347 snThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 367
Db 858 ----- 858
Qy 367 ysLeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsnPro**S 387
Db 858 ----- 858
Qy 387 erAlaMetAlaArgProLeu**ArgProArgArgAspIleTyrGluAspIleHisAlaT 407
Db 859 -----NNNNNGATATCTGGGAAGACGTACATGCTT 889
Qy 407 rp 407
Db 890 GG 891
RESULT 12
MMPAR2MR
LOCUS M.musculus mRNA for proteinase activated receptor 2.
DEFINITION 2713 bp mRNA linear ROD 08-JUN-1995
ACCESSION 248043
VERSION 248043.1 GI:663020
KEYWORDS proteinase activated receptor 2.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2713)
Nystedt, S., Larsson, A.K., Aberg, H. and Sundelin, J.
The mouse proteinase-activated receptor-2 cDNA and gene. Molecular
cloning and functional expression
J. Biol. Chem. 270 (11), 5950-5955 (1995)
MEDLINE 95197620
PUBMED 7890726
REFERENCE 2 (bases 1 to 2713)
AUTHORS Nystedt, S.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-1995) Sverker Nystedt, Division of Molecular
Neurobiology, The Wallenberg Laboratory, Lund University,
Soelvegatan 33A, Lund, S-220 07, Sweden
FEATURES
Location/Qualifiers
source 1..2713
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="cm140pBS"
5'UTR 1..72
/citation=[1]
CDS 73..1272
/codon_start=1
/codon_usage=[1]
/product="proteinase activated receptor 2"
/protein_id="CAA88097.1"
/db_xref="GI:663021"
/db_xref="MGD:MGI:101910"
/db_xref="SWISS-PROT:P55086"
/translation="MRSLSLAWLGGITLLAASVSCRSRTENLAPGRNNSKGRSLIGRL
ETQPPITKGVPVEPGFSIDEPSASILTKLATVFLPVVYIIVFVIGLPSNGMALWIF
LRTKKKHPAVLYMANLALDELSSLWIFPKISYHLHGNWVYGEALCKVLIGFFYGN
MYCSILFMTCLSVQRYWVIVNPMGHPKKANAVGSLAILLILFLVTLPLYVMKQTI
YIPALNITTCDBVLPEEVLVGMFNFLSLAIGVFLPALLTASAYVLMIKTLRSSAM
DHSEKKRQRAIRLIITVLNMYFICFAPSNLLVHVHYFLIKTORQSHVVALYVALUCL
STNSCIDPFVYFVSKDFRHNALLCRSVRTVMRQIISLSSNFKFSRSGSYSSSS
TSVKTSY"
intron 163..164
/notes="genomic intron location"
3'UTR
/citation=[1]
1273..2713
/polyA_signal 2219..2224
/citation=[1]
polyA_signal 2693..2698
BASE COUNT 632 a 676 c 633 g 772 t
ORIGIN
Alignment Scores:
Pred. No.: 3..59e-38 Length: 2713
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 10 Gaps: 10
US-09-208-629f-3 (1-407) x MMPAR2MR (1-2713)
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db 61 CCACGTCCGGGATGCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyLeuAsnValSer-----Asp 43
Db 112 ACCCTTCTGGCGCCTCTGGTCTCTCTGACCGGACCGAAGACCTTGACCGGACCGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 AACAGTAAAGGAAGAAGTCTTATTGGCAGATTAGAAAACCCAGCTCCAATCACTGGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACGAGCTTTTCCATGATGAGTTCTCTCGCTCCATCTCACCGGG 291
Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db ----- 113
```


Db 790 GACATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCCCGGCCCTCCTT 849
 Qy 293 IleilePheCysTyrThrLeuLeu-----His 302
 Db 850 ACTGCATCTGCTAGTCTCATGATCAAGACGCTCGCTCTTCTGTATGGATGAACAC 909
 Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
 Db 910 TCAGAGAGAAAGCGAGAGGCT-----ATCCGACTCATCATCACCCTGCTG 957
 Qy 323 ValilePheThrileCysPheAlaProThrAsnIleleLeuValileHisAlaAsn 342
 Db 958 GCCATGTACTCATCTGCTTGTCTCTAGCAACCTTCTGTCTGTAGTGCATTATTTCCTA 1017
 Qy 343 TyrTyrThrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
 Db 1018 ATCAAAACCCAGAGAGAGCCAGCTCTACGCCCTCTACCTTGTGCGCCCTCTGCTCTG 1077
 Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCAAAA 1125

 RESULT 14
 AR171259 2732 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 60 from patent US 6297026.
 ACCESSION AR171259
 VERSION AR171259.1 GI:17910209
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2732)
 AUTHORS Sundelin,J. and Scarborough,R.M.
 TITLE Nucleic acids encoding the C140 receptor
 JOURNAL Patent: US 6297026-A 60 02-OCT-2001;
 FEATURES Location/Qualifiers
 source 1..2732
 BASE COUNT 650 a 676 c 633 g 773 t
 ORIGIN

 Alignment Scores:
 Pred. No.: 3.62e-38 Length: 2732
 Score: 584.00 Matches: 128
 Percent Similarity: 51.52% Conservative: 76
 Best Local Similarity: 32.32% Mismatches: 126
 Query Match: 27.34% Indels: 66
 DB: 6 Gaps: 10

 US-09-208-629F-3 (1-407) x AR171259 (1-2732)
 Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
 Db 61 CCAGCTCCGGGATCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111
 Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
 Db 112 ACCCTTCTGGCGCTCGGTCTCTCTGAGCCGAGAACCTTGCACCGGAGCGCAAC 171
 Qy 44 AsnSer-----AlaLysProThrLeuThrileLys 53
 Db 172 AACAGTAAAGGAGAGAGTCTTATTGGCAGATTAGAAACCCAGCTCCAATCCTGGGAA 231
 Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIleGluGly 73
 Db 232 GGGGTTCCGGTAGAACAGGCTTTTCCATCGATGAGTCTCTGCGCTCCATCCTCACC 291
 Qy 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
 Db 291 ----- 291
 Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnVallePro 113

Db 292 -----AAGCTGACCACGGTCTTCTTCCTCG 315
 Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
 Db 316 GTCTGTACATTAATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
 Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
 Db 376 TTCTTTTTCGAGACGAGAGAAACACCCCGCGTGTATTTACATGGCCAACTTGGCCCTG 435
 Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
 Db 436 GCCGACCT 495
 Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
 Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 555
 Qy 193 TyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
 Db 556 TATTGTCTCATCT 615
 Qy 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
 Db 616 CCCATGGGACACCCCGAG---AAGAAGGCAACATCGCGTGTGGCGTCTCTCTTGGCAATC 672
 Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
 Db 673 TGGCTCTCTGATTTTCTGTGTCACCATCCCTTTGTATGTCTATGAAGCAGACCATCTACATT 732
 Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
 Db 733 CCAGCATTTGACATCACCACCTGTACGATGTCTGCTGCTGAG---GAGGTATTGGTGGG 789
 Qy 273 PheArgPheTyrThrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
 Db 790 GACATGTTCAATTACTTCT 849
 Qy 293 IleilePheCysTyrThrThrLeuLeu-----His 302
 Db 850 ACTGCATCTGCTACGCTCATGATCAAGACGCTCGCTCTTCTGTCTATGGATGAACAC 909
 Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
 Db 910 TCAGAGAGAAAGCGAGAGGCT-----ATCCGACTCATCATCACCCTGCTG 957
 Qy 323 ValilePheThrileCysPheAlaProThrAsnIleleLeuValileHisAlaAsn 342
 Db 958 GCCATGTACTCATCTGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
 Qy 343 TyrTyrThrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
 Db 1018 ATCAAAACCCAGAGAGAGCCAGCTCTACGCCCTCTACCTTGTGCGCCCTCTGCTCTG 1077
 Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCAAAA 1125

 RESULT 15
 187850
 LOCUS 187850 2732 bp DNA linear PAT 10-AUG-1998
 DEFINITION Sequence 60 from patent US 5716789.
 ACCESSION 187850
 VERSION 187850.1 GI:3407790
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2732)
 AUTHORS Sundelin,J. and Scarborough,R.M.
 TITLE Method to determine ligands, agonist and antagonist of C140 receptor

JOURNAL Patent: US 5716789-A 60 10-FEB-1998;

FEATURES

Location/Qualifiers
1. .2732

BASE COUNT 650 a 676 c 633 g 773 t

ORIGIN

Alignment Scores:

Pred. No.: 3.62e-38 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservatives: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 6 Gaps: 10

US-09-208-629F-3 (1-407) x 187850 (1-2732)

Qy 8 ProValAlaGlySerGlnAspIleLeuMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db 61 CCAGTCCGGGATGCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db 112 ACCCTTCTGGCGGCTGGTCTCTCGAGCGGACCGAAGACCTTGACCGGAGCGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 AACAGTAAGGAAGAGTCTATTGGCAGATTAGAAACCCAGCTCCATCACTGGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACACGAGGCTTTTCCATCATGATGAGTTCTCTGGCTCCATCTCACC 291
Qy 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerLeu 93
Db 291 -----
Qy 94 HisValAsnAsnAlaThrIleGlyTyLeuArgSerSerLeuSerThrGlnValIlePro 113
Db 292 -----AAGCTGACCAACCGGCTTCTTCTCCG 315
Qy 114 AlaIleTyIleLeuPheValGlyValProSerAsnIleValThrLeuTrpLys 133
Db 316 GTCGTCTACATATTGTGTTGTGATGTTTGGCCAGTAATGCGCCCTCTGCGCTCGATC 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCCTTTCCGAAGAAAGAAAGAACACCCCGCGGTGATTTACATGGCCACCTGGCCTTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsn 172
Db 436 GCGGACCTCTCTCTGTCATCTGGTTCCTCCCTGAAGATCTCTTACCACTACATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCGCTGTGAAGTGTCTATGGCTTTTCTATGTTAATGTAACATG 555
Qy 193 TyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyLeuAlaThrAlaHis 212
Db 556 TATTGCTCCATCCCTCTCATGACCTGCTGAGTGTGAGGAGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCCATGGACACCCAGG---AAGAGGCAACATCGCGTGGCTCTCTTGGCAATC 672
Qy 233 TrpValMetValPheLeuTyMetLeuProPheValIleLeuLysGlnTyHisLeu 252
Db 673 TGGCTCCTGATTTTCTGGTCACCATCTTGTGATGTCATGAAGCAGCACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATCACACACCTGTCAGATGTCGCTGAG---GAGGTATTGGTGGG 789

Qy 273 PheArgPheTyTrpPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTTCATTAATTACTTCTCTCACTGGGCAATTGGAGTCTTCTGTTCGGGCTCCT 849
Qy 293 IleIlePheCysTyTrpThrThrLeuIle-----His 302
Db 850 ACTGCATCTCGCTACGTGCTCATGATCAAGACCTCGCTCTTCTGTATGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyIleLysAlaValLeuLeuLeu 322
Db 910 TCAGAGAGAGAAAGGAGGAGGCT-----ATCGACTCATCATCCGCTGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsn 342
Db 958 GCCATGTACTTCTATCTGCTTGTCTAGCAACCTTCTGCTAGTGCATTATTTCCTA 1017
Qy 343 TyTrpTyTrpHisAsnThrAspSerLeuTyPheMetTyLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCCAGAGGAGGAGGAGGCTTACGCTTCTACCTTGTGCGCTCTGCTGCTG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyTrpPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTGTCTCAAAA 1125
RESULT 16
BC025432 2772 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, coagulation factor II (thrombin) receptor-like 1,
DEFINITION clone MGC:29183 IMAGE:5006789, mRNA, complete cds.
ACCESSION BC025432
VERSION BC025432.1 GI:19343971
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Strausberg, R.
Direct Submission
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Breen, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 39 Row: d Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
FEATURES
source
1. .2772
/organism="Mus musculus"
/db_xref="taxon:10090"

```

/map="FVB/N-3"
/clone="MGC:29183 IMAGE:5006769"
/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
old mouse, Taken by biopsy."
/clone_lib="NCI CGAP Mam2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
99..1298
/codon_start=1
/product="coagulation factor II (thrombin) receptor-like
1"
/protein_id="AAH25432.1"
/db_xref="GI:19343972"
/db_xref="LocusID:14063"
/translation="MRSLSLAWLLGGITLAAVSVCSTENLAPGRNNSKGRSLIGRL
ETQPIITGKGVPEPFGSDEFASILITGKLTIVLPVYIIIVFVIGLPSNGMALWIF
LFTKKHAPAVIYMANIALADLSVIFPLAIYAHGLGNWVYGEALCKVLIGFFYGN
MYGSLFMTCLSVQRYWVIVPMGHPKRNIAVGVSLAIWLIIFLIVTPLYVMKQTI
YIPALNITCTHDVLPPELVGDMFNFLSLAIGVFLPAILTASAVVLMKLTURSLM
DEHSEKKRQRAIRLIITVLAMYRICFAPSRLLVVHFLIKTORQSHVVALYVALCL
STLNSCIDPFPYFVSVKDFRDHARNALLCRSVTRNMQISLNSKFSKSGSYSSSS
TSVKTSY"
BASE COUNT 671 a 689 c 638 g 774 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 4,44e-38 Length: 2772
Score: 583.00 Matches: 128
Percent Similarity: 51.26% Conservative: 75
Best Local Similarity: 32.32% Mismatches: 127
Query Match: 27.29% Indels: 66
DB: 10 Gaps: 10

```

US-09-208-629F-3 (1-407) x BC025432 (1-2772)

```

Qy 8 ProValAlaGlySerGlnAspIleLeuMetIleLeuValAlaGlyLeu 27
Dy 87 CCAGTCCGGGATGCGAGTCTCAGCTGCGGTGCTG-----GGAGGTATC 137
Qy 28 LeuPheLeuProValThrVal-----CysGlnSerGlyIleAsnValSer-----Asp 43
Dy 138 ACCCTTCTGCGGCGCTCGGTCTCTCGAGCGGACCGAGAACCTTGCACCGGAGCAAC 197
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Dy 198 AACAGTAAAGGAAGTCTATTGGCAGATTGGAACCCAGCTCCATCCTCAGTGGAAA 257
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIleGluGly 73
Dy 258 GGGGTTCGGTAGAACACGAGCTTTCCATCGATGAGTCTCTCGCTCCATCCTCACCGGG 317
Qy 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Dy 317 ----- 317
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Dy 318 -----AAGTGCACCATCTGCTTTCTTCGG 341
Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLys 133
Dy 342 GTCGTCTACATTATGTTGTGATGTTTCCAGTAATGATGATGCGCTCTCGATC 401
Qy 134 LeuSerLeuArgThr-----LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Dy 402 TTCCTTTTCCGAACGAAGAAACACCCCGCGGTGATTTACATGGCCAACTCGGCCTTG 461
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Dy 462 GCTGACCTCTCTCTGTCATCTGGTTCCTCCCTGGCCATGCTACCACTACATGGCAAC 521
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValPheTyrGlyAsnMet 192
Dy ----- 192

```

```

Db 522 AACTGGGTCTATGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGCAACATG 581
Qy 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaHis 212
Db 582 TATTGCTCATCTCTTTCATGACCTCCCTCAGGTGCAGAGTACTGGGTGATCGTGAAC 641
Qy 213 PropheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 642 CCCATGGGACCCCAAGAGGCAACATCGCGTGGCTCTCTCTGGCAATC 698
Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnLysHisLeu 252
Db 699 TGGCTCCTGATTTTCTGTCACCATCTCTTGTATGTCATGAAGCAGACCATCTACAT 758
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 759 CCAGCTTGAACATCACCATCTGCCAGATGCTCCCTGAG---GAGGTATTGGTGGGG 815
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 816 GACATGTTCAATTACTTCTCTCAGTGGCATTTGGAGTCTTCTGTTCCGGCCATCCT 875
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 876 ACTGATCTCGCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGATGATGAACAC 935
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db 936 TCAGAGAAGAAAGGAGGAGGCT-----ATCCGATCATCATCACCGTGTG 983
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 984 GCCATGTTCTATCTGCTTGTCTCTAGCACTTCTGCTGATGATGATTTCTCCTA 1043
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGly 362
Db 1044 ATCAAAACCCAGAGGAGGAGCAGCCAGCTACGCCCTCTACCTGCTGCGCTCTGCTG 1103
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1104 ACTCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCTCAAAA 1151

RESULT 17
XLU09632 1680 bp mRNA linear VRT 28-MAY-1994
LOCUS Xenopus laevis thrombin receptor mRNA, complete cds.
DEFINITION U09632
ACCESSION U09632.1 GI:495197
VERSION 1
KEYWORDS African clawed frog.
SOURCE Xenopus laevis
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 420)
AUTHORS Gersten,R.E., Chen,J., Ishii,M., Ishii,K., Wang,L., Nanevicz,T.,
Turck,C.W., Vu,T.K. and Coughlin,S.R.
TITLE Specificity of the thrombin receptor for agonist peptide is defined
by its extracellular surface
JOURNAL Nature 368 (6472), 648-651 (1994)
MEDLINE 94195429
PUBMED 8145852
REFERENCE 2 (bases 1 to 1680)
AUTHORS Coughlin,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1994) Shaun R. Coughlin, CVRI, University of
California, San Francisco, 505 Parnassus Avenue, Room HSW 831, San
Francisco, CA 94143, USA
FEATURES
source Location/Qualifiers
1..1680 /organism="Xenopus laevis"
/db_xref="taxon:8355"
197..1459
CDS

```

```

/codon_start=1
/products="thrombin receptor"
/protein_id="AAA18498.1"
/db_xref="GI:495198"
/translations="MMELRVLLLLLLLLTLAGMSLCLANSIDTOAGAHNNMTIKTF
RIFDSESEFEIIPWDELDESGSGDOAPVRSARKPIRNITKEAQYLSQWLTK
FVSLYTVFVIGLPLNLLAIILFLPKVRKPAVYMLNLAIDVSVLSFLPKIAY
HLSGNLWLFPGMCRIVTAIFYCNMYCSVLLIASISVDRELVAVYPMHLSWETMERA
YMACSFILWISASTIPPLVTEQTKIPRLDITCHDLVDLKDQDFIYIFYSFCLL
FFVPFIIITICVIGIIRSLSSSIENSCKKTRFLAVVLVCVFIICFQTNVLFIT
HYLQENEFYFAVILSACVGSCLDPLIYIYASSQCCQYLYLSLLCCRKVEPSSS
TGQLMSPAMKNDNCSTNAXSIYKLLA"
BASE COUNT 427 a 366 c 401 g 486 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 2,78e-38 Length: 1680
Score: 582.50 Matches: 124
Percent Similarity: 54.35% Conservativeness: 82
Best Local Similarity: 32.72% Mismatches: 138
Query Match: 27.27% Indels: 35
DB: 5 Gaps: 11

```

US-09-208-629F-3 (1-407) x XLU09632 (1-1680)

```

QY 19 IleuIleuLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGly 38
DB 215 CTGCTGCTGCTGCTGCTATTGACCTCTGGGAGCGATGGGATCTCTCTCTGGCCAA 274
QY 39 IleAsnValSerAspAsnSerAlaLys-----ProThrLeuThrIleLysSer 54
DB 275 -----TCTGATACCCAGCAAGAGGGTGCTCATTTCAACAATATGACCATATAAAAT 325
QY 55 PheAsn-----GlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGlu 72
DB 326 TTGAGATATTTGACGACAGTGAAGTGGTGAAGAGATTCATCGGATGAACATCGAT 385
QY 73 GlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThr 92
DB 386 GAA---TCAGGAGAGAGTTCCAGGAGACCGCCCGGTGTCCTAGTGAAGAGGCCA 442
QY 93 Leu-----HisValAsnAsnAlaThrIleGlyTyrlleuArgSerLeuSerThrGln 110
DB 443 ATCCGGAGGAAACATCACAAAGGAGCGGAGCAATATCTCTCCAGTCTAGTGGTGA 502
QY 111 ValIleProAlaIleTyrlleLeuLeuPheValValGlyValProSerAsn----- 127
DB 503 TTGTGACATCCCTGTACATGCTGCTTTATTTGTTGGGCTGCTCTGAAATCTGTGGCA 562
QY 128 ---IleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePhe 146
DB 563 ATCATCATATTTCTGTTCAAGATGAAAGTCAGAAAACCGGCA-----GTGGTCTAT 613
QY 147 HisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAla 166
DB 614 ATGTGTAACCTGGCCATCGCCGACGCTGCTTTGTTAGTGTGTGCTTTCAAAATAGCC 673
QY 167 TyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrVal 186
DB 674 TACCATCTGCGGAAACGAGTGGCTGTTGGGCTGGATGTGCCCATGTGCATGCTGCC 733
QY 187 ValPheTyrlleGlyAsnMetTyrlleCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArg 206
DB 734 ATCTTCTACTGCAACATGTACTGCTGCTGCTGCTCATCGCCAGCATCAGTGTGGACAGG 793
QY 207 TyrlleuAlaThrAlaHisProPheThrTyrlleGlnLysLeuProLysArgSerPheSerLeu 226
DB 794 TTCTTGGCAGTGTTTATTCACCATGCACTCCCTTTCTTTGGGCGCAACATGAGCCGTGCTAT 853
QY 227 LeuMetCysGlyIleValTrpValMetValPheLeuTyrlleMetLeuProPheValIleLeu 246
DB 854 ATGGCCTGTAGTTTCATTTGGCTGATATCCATGAGCGACATATACCATCTTCTGTGACT 913

```

```

QY 247 LysGlnGluTyrlleHisLeuValHisSerGluIleThrThrCysHisAspValValAspAla 266
DB 914 GAGCAAACTCAGAAAGATTCTAGCTGGATATTACTACTTCCATGATGTTTAGAT--- 970
QY 267 CysGluSerProSerSerPheArgPheTyrllePheValSerLeuAlaPhePheGlyPhe 286
DB 971 TTGAAAGACTCMAAGACTTTATATCTACTATTCTTCTTCTTCTTCTTCTTCTTCTTCT 1030
QY 287 LeuIleProPheValIleIleIlePheCysTyrlleThrThrLeuIleHisLysLeuLysSer 306
DB 1031 TTGCTGCCATTCATTATCACCACATCTCTCTACATAGGAGATCATCAGAAGCTTAAGTTCC 1090
QY 307 -----LysAspArgIleTrpLeuGlyTyrlleLysAlaVal 318
DB 1091 TCCAGCATCGAGAACAGTTGCAAGAAAGACAGAGCC-----CTGTCTTCTGGCG 1138
QY 319 LeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIle 338
DB 1139 GTGGTTGTGCTGCTGCTGCTTTTATCATCTGTTTGGCCCAACCAATGCTCTTCTTCTGACC 1198
QY 339 HisHisAlaAsnTyrlleTyrlleHisAsnThrAspSerLeuTyrlleMetTyrlleAla 358
DB 1199 CAT-----TACTTGAAGAGAGCCAAATGATTTCTATATTTGCTTACATTCGAGT 1249
QY 359 LeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrlleValMetSer 377
DB 1250 GCCTGTGTTGGCAGCGTCAGCTGTGCTCGATCCCTTATTTACTACTATGCGTCA 1306
RESULT 18
MMPROACR2 1477 bp DNA linear ROD 30-JAN-1995
LOCUS M.musculus exon 2 for proteinase activated receptor 2.
ACCESSION Z35158.1 GI:559918
VERSION 235158.1
KEYWORDS proteinase activated receptor; proteinase activated receptor 2.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1477)
AUTHORS Nystedt,S., Emilsson,K., Wahlestedt,C. and Sundelin,J.
TITLE Molecular cloning of a potential proteinase activated receptor
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (20), 9208-9212 (1994)
MEDLINE 95023880
PUBMED 7937743
REFERENCE 2 (bases 1 to 1477)
AUTHORS Nystedt,S.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1994) Sverker Nystedt, Division of Neurobiology,
The Wallenberg, Laboratory, Lund University, Soelvegatan 33A, Lund,
S-220 07, Sweden
REFERENCE 3 (bases 1 to 1477)
AUTHORS Nystedt,S., Larsson,A.K., Aaberg,H. and Sundelin,J.
TITLE The mouse proteinase activated receptor 2 cDNA and gene-molecular
cloning and functional expression
JOURNAL J. Biol. Chem. 170 (1995) In press
FEATURES
Location/Qualifiers
source 1..1477
/organism="Mus musculus"
/db_xref="taxon:10090"
/germline
gene 1..1477
/gene="PAR-2"
intron <1..310
/gene="PAR-2"
/citation=[3]
exon 311..1477
/gene="PAR-2"
/citation=[3]
/number=1
1420..1477
/gene="PAR-2"
3'UTR 1420..1477
BASE COUNT 341 a 413 c 321 g 402 t

```

```

AC110876/c
LOCUS      199227 bp      DNA      linear      HTG 16-FEB-2002
DEFINITION Mus musculus clone RP23-176N6, WORKING DRAFT SEQUENCE, 8 unordered
           pieces.
AC110876
AC110876.1 GI:18693418
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 199227)
AUTHORS     Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 199227)
AUTHORS     Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and
            Cheng,J.-F.
TITLE       Direct Submission
JOURNAL     Submitted (16-FEB-2002) Berkeley PGA, Lawrence Berkeley National
            Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
            America
COMMENT     Draft Sequence Produced by Berkeley PGA
            pga.lbl.gov
            Center Code: PGABERK.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            * 1 2263: contig of 2263 bp in length
            * 2264 2363: gap of unknown length
            * 2364 5169: contig of 2806 bp in length
            * 5170 5269: gap of unknown length
            * 5270 13742: contig of 8473 bp in length
            * 13743 13842: gap of unknown length
            * 13843 21596: contig of 7754 bp in length
            * 21597 21696: gap of unknown length
            * 21697 37430: contig of 15734 bp in length
            * 37431 37530: gap of unknown length
            * 37531 77278: contig of 39748 bp in length
            * 77279 77378: gap of unknown length
            * 77379 129890: contig of 52512 bp in length
            * 129891 129991: gap of unknown length
            * 129991 199227: contig of 69237 bp in length.
FEATURES             Location/Qualifiers
     source           1..199227
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT  55269 a 45000 c 45562 g 52676 t 720 others
ORIGIN

Alignment Scores:
Pred. No.:      7,07e-36      Length:      199227
Score:          581.50        Matches:      111
Percent Similarity: 59.66%    Conservative: 65
Best Local Similarity: 37.63% Mismatches:    102
Query Match:    27.22%       Indels:       17
DB:             Gaps:        5

US-09-208-629F-3 (1-407) x MMPROACR2 (1-1477)
Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTCTCTGCGTCATCTCACCGGAGCTGACCAGGTCCTTCTTCGCGTC 465
Qy 115 IleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTrpLysLeu 134
Db 466 GTCTACATTATTGTGTTGTGATTGTTGCCAGTAATGCGATGCCCTCTGATCTTC 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAAGAGAAAGAAACACCCCGCGTGAATTCATGGCCAACTGGCGCTGGCC 585
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTGTCATCTGGTCCCTCGAGATCTCTACCACCTACATGGCAACAC 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGTGCTCATTTGGCTTTTCTATGTAACATGTAT 705
Qy 194 CysAlaIleLeuLeuLeuTyrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTTCATGACCTGCTCAGCGTGAGAGGACTGGGTGATCGTGAACCCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
Db 766 ATGGGACACCCAGG---AAGAAGGCAACATCCCGTGGCGTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 823 CTCCTGATTTTCTGGTCACCATCTTGTATGTCATGAAGCAGACCATCTACATTTCCA 882
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273
Db 883 GCATTGAACATCACACCTGTACAGATGTGCTGCTGAG---GAGGTATTGGTGGGGGAC 939
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293
Db 940 ATGTTCAATTAATCTCTCTCACTGGCCATGGAGTCTTCTGTTCGCGCCCTCTTACT 999
Qy 294 IlePheCysTyrThrThrLeuIle-----HisLys 303
Db 1000 GCATCTGCGCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTGATGGAACACTCA 1059
Qy 304 LeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuVal 323
Db 1060 GAGAAGAAAGCGAGAGGGCT-----ATCCGACTCATCATCAGCTCGCTCGGCC 1107
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyr 343
Db 1108 ATGTACTTCATCTGCTTGTCTCTAGCAACCTTCTGCTGAGTGCAATTATTCTCTAATC 1167
Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleIleAlaLeuCysLeuGlySer 363
Db 1168 AAAACCCAGAGGAGCAGCAGCTGTACGCCCTCTACTTGTTCGCGCTCTGCTCGTGCAC 1227
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1228 CTCACAGCTGCATGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 19

```

Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
| | | | |
Db 174361 CTTTTCGACGAAGAAACACCCCGCGTGTATTTACATGGCCAACTGGCGCTGGCC 174302
| | | | |
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrrHisLeuAsnGlyAsnAsn 173
| | | | |
Db 174301 GACCTCTCTCTGTCTGTGGTCCCTCGAAGATCTCCACCACTACATGGCAACAAC 174242
| | | | |
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrrGlyAsnMetTyr 193
| | | | |
Db 174241 TGGGTCTACGGGAGGCCCTGTGCAAGGTCTCATTTGGCTTTTCTATGGTAACATGTAT 174182
| | | | |
Qy 194 CysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrrLeuAlaThrAlaHisPro 213
| | | | |
Db 174181 TGTCTCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 174122
| | | | |
Qy 214 PheThrTyrrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
| | | | |
Db 174121 ATGGGACACCCCGAGG---AAGAAGGCAACATCCCGGTGGCGTCTCTTGGCAATCTGG 174065
| | | | |
Qy 234 ValMetValPheLeuTyrrMetLeuProPheValIleLeuLysGlnGluTyrrHisLeuVal 253
| | | | |
Db 174064 CTCTGATTTTCTGGTTCACCATCCCTTTGTATGTCTATGAGCAGACCATCTACATTTCCA 174005
| | | | |
Qy 254 HisSerGluLeuThrThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273
| | | | |
Db 174004 GCATTTGAACATCACCATCTGTACCATGTGTCTGCTGAG---GAGGTATTGGTGGGGGAC 173948
| | | | |
Qy 274 ArgPheTyrrTyrrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleIle 293
| | | | |
Db 173947 ATGTTCAATTTACTTCTCTCAGTGGCCATTTGGAGTCTTCTGTTTCCCGGCCCTCTTACT 173888
| | | | |
Qy 294 IlePheCysTyrrThrThrLeuIle-----HisLys 303
| | | | |
Db 173887 GCATCTCGCTACGGTCTCATGATCAAGCGCTCGCTCTCTCTGCTATGGATGAACATCA 173828
| | | | |
Qy 304 LeuLysSerIleAspArgIleTrpLeuGlyTyrrIleLysAlaValLeuLeuLeuVal 323
| | | | |
Db 173827 GAGAAGAAAGGAGGAGGGCT-----ATCGGATCATCATCATCGTGTGGCC 173780
| | | | |
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
| | | | |
Db 173779 ATGTACTTCTATCTGCTTTGTCTCTAGCAACCTCTCTGCTGAGTGCATTTATTTCTTAATC 173720
| | | | |
Qy 344 TyrrTyrrHisAsnThrAspSerLeuTyrrPheMetTyrrLeuIleAlaLeuCysLeuGlySer 363
| | | | |
Db 173719 AAAACCCAGAGGAGAGCCAGCTCTACGCCCTCTACCTTTGTCGCCCTCTGCTGTGACCC 173660
| | | | |
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrrPheValMetSerTyr 378
| | | | |
Db 173659 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 173615
| | | | |

RESULT 20

AR012637 LOCUS 1475 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5763575.
ACCESSION AR012637
VERSION AR012637.1 GI:3970627
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin, J. and Scarborough, R. M.
TITLE Agonist and antagonist peptides of the C140 receptor
JOURNAL Patent: US 5763575-A 1 09-JUN-1998;
FEATURES Location/Qualifiers
source 1. 1475
/organism="unknown"

BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN

Alignment Scores:

Pred. No.: 4,21e-38 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: Gaps: 5
US-09-208-629F-3 (1-407) x AR012637 (1-1475)
Qy 95 ValAsnAsnAlaThrIleGlyTyrrLeuArgSerSerLeuSerThrGlnValIleProAla 114
| | | | |
Db 406 ATCGATGAGTTCTTCGCTCCATCTCCACCGGAAGCTGACCAAGTCTTTTCCGGTC 465
| | | | |
Qy 115 IleTyrrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeu 134
| | | | |
Db 466 GTCTACATTATTCTGTTTGTGATTGTTTGCCTCAGTAATGGCATGGCCCTCTGGATCTTC 525
| | | | |
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
| | | | |
Db 526 CTTTTCGACGAAGAAAGAAACACCCCGCGTGTATTTACATGGCCAACTGGCGCTGGCC 585
| | | | |
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrrHisLeuAsnGlyAsnAsn 173
| | | | |
Db 586 GACCTCTCTCTGTCTCATCTGCTGTTCCCTCGAAGATCTCTTACCACCTTACATGGCAACAAC 645
| | | | |
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrrGlyAsnMetTyr 193
| | | | |
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGGTCTCATTTGGCTTTTCTATGGTAACATGTAT 705
| | | | |
Qy 194 CysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrrLeuAlaThrAlaHisPro 213
| | | | |
Db 706 TGTCTCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 765
| | | | |
Qy 214 PheThrTyrrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
| | | | |
Db 766 ATGGGACACCCCGAGG---AAGAAGGCAACATCCCGGTGGCGTCTCTTGGCAATCTGG 822
| | | | |
Qy 234 ValMetValPheLeuTyrrMetLeuProPheValIleLeuLysGlnGluTyrrHisLeuVal 253
| | | | |
Db 823 CTCTGATTTTCTGGTCAACATCCCTTTGTATGTCTATGAGCAGACCATCTACATTTCCA 882
| | | | |
Qy 254 HisSerGluLeuThrThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273
| | | | |
Db 883 GCATTTGAACATCACCATCTGTACCATGTGTCTGCTGAG---GAGGTATTGGTGGGGGAC 939
| | | | |
Qy 274 ArgPheTyrrTyrrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleIle 293
| | | | |
Db 940 ATGTTCAATTTACTTCTCTCAGTGGCCATTTGGAGTCTTCTGTTCCCGGCCCTCTTACT 999
| | | | |
Qy 294 IlePheCysTyrrThrThrLeuIleHisLysLeuLysSer----- 306
| | | | |
Db 1000 GCATCTGCTCTACGTGCTCATGATCAAGCGCTCGCTCTTCTGCTATGGATGAACATCA 1059
| | | | |
Qy 307 -----LysAspArgIleTrpLeuGlyTyrrIleLysAlaValLeuLeuLeuVal 323
| | | | |
Db 1060 GAGAACAAGGAGGAGGGCT-----ATCGGATCATCATCATCGCTGTGGCC 1107
| | | | |
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
| | | | |
Db 1108 ATGTACTTCTATCTGCTTTGTCTCTAGCAACCTTCTGCTGCTAGTGCATTTATTTCTTAATC 1167
| | | | |
Qy 344 TyrrTyrrHisAsnThrAspSerLeuTyrrPheMetTyrrLeuIleAlaLeuCysLeuGlySer 363
| | | | |
Db 1168 AAAACCCAGAGGAGGAGCCAGCTCTACGCCCTCTACCTTTGTCGCCCTCTGCTGTGTCGACC 1227
| | | | |
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrrPheValMetSerTyr 378
| | | | |
Db 1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272
| | | | |

RESULT 21

AR171257 LOCUS 1475 bp DNA linear PAT 17-DEC-2001

```
DEFINITION Sequence 1 from patent US 6297026.
ACCESSION AR171257
VERSION AR171257.1 GI:17910207
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Nucleic acids encoding the C140 receptor
JOURNAL Patent: US 6297026-A 1 02-OCT-2001;
FEATURES
    Location/Qualifiers
        1..1475
            /organism="unknown"
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN
Alignment Scores:
Pred. No.: 4.21e-38 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: 6 Gaps: 5

US-09-208-629F-3 (1-407) x AR171257 (1-1475)
Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTTCTCTGGTCCATCTCACCAGGAACTGACCGGTCCTTCTTCCTCCGTC 465
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAACGAGAAAGAACACCCCGCGTGATTACATGCGCAACCTTGGCCTTGGCC 585
Qy 154 AspleuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 705
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCCGAGG---AAGAAGGCAACATCGCCGTGGCGTCTCTCTCTCTCTCTCT 822
Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253

US-09-208-629F-3 (1-407) x I42454 (1-1475)
Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTTCTCTGGTCCATCTCACCAGGAACTGACCGGTCCTTCTTCCTCCGTC 465
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAACGAGAAAGAACACCCCGCGTGATTACATGCGCAACCTTGGCCTTGGCC 585
Qy 154 AspleuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 705
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCCGAGG---AAGAAGGCAACATCGCCGTGGCGTCTCTCTCTCTCTCTCT 822
Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
```



```
Db      823 CTCCTGATTTTCTGGTCCACATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCA 882
Qy      254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273
Db      883 GCATTGAACATCACCATCTGTCAGATGCTGCTGAG--GAGGTATTGGTGGGGGAC 939
Qy      274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db      940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCCCGGCCCTCTTACT 999
Qy      294 IlePheCysTyrThrThrLeuIleHisLeuLysSer----- 306
Db      1000 GCATCTGCCGTACGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059
Qy      307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
Db      1060 GAGACAAAGGCAGAGGGCT-----ATCGACTCATCATCAGCGTGGTGGCC 1107
Qy      324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
Db      1108 ATGACTTCACTCTGCTTGTCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATC 1167
Qy      344 TyrThrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGlySer 363
Db      1168 AAAACCCAGGCGAGCCAGTCTACGCCCTCTACCTTGTGCGCCCTCTGCTGTCGACC 1227
Qy      364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db      1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 23
LOCUS      I87848      1475 bp      DNA      linear      PAT 10-AUG-1998
DEFINITION Sequence 1 from patent US 5716789.
ACCESSION  I87848
VERSION    I87848.1 GI:3407788
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1475)
AUTHORS    Sundelin,J. and Scarborough,R.M.
TITLE      Method to determine ligands, agonist and antagonist of C140
           receptor
JOURNAL    Patent: US 5716789-A 1 10-FEB-1998;
FEATURES   Location/Qualifiers
           source
           1..1475
           /organism="unknown"
BASE COUNT 341 a 414 c 319 g 401 t

Alignment Scores:
Pred. No.: 4.21e-38      Length: 1475
Score: 579.50      Matches: 111
Percent Similarity: 60.00%      Conservative: 66
Best Local Similarity: 37.63%      Mismatches: 101
Query Match: 27.13%      Indels: 17
DB: 6      Gaps: 5

US-09-208-629F-3 (1-407) x I87848 (1-1475)

Qy      95 ValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db      406 ATCATGAGTCTCTGGGTCTCATCTTACCGGAGCTGACCGGTCTTCTTCCGGTC 465
Qy      115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeu 134
Db      466 GTCTACATTATTGTGTTGTGATTGTTTGGCCAGTAATGCGATGGCCCTCTGGATCTTC 525
Qy      135 SerLeuArgThr-----LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db      526 CTTTTCGACGAGGAGAAACACCCCGCGGTGATTTTACATGGCCAACTGGCCTTGCC 585
```

```
Qy      154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db      586 GACCTCTCTCTCTGTCATCTGGTTCCCTCGAGAGATCTCTTACCACCTACATGCAACAAC 645
Qy      174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db      646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGTAT 705
Qy      194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db      706 TGTCTCATCTCTTTCATGACCTGCTCAGGTCGACAGGTACTGGGTGATCGTGAACCCC 765
Qy      214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
Db      766 ATGGGACACCCACAG--AAGAAGGCAACATCGCGTGGCTCTCTTGGCAATCTGG 822
Qy      234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db      823 CTCCTGATTTTCTGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCA 882
Qy      254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db      883 GCATTGAACATCACCATCTGTCAGATGCTGCTGCTGAG--GAGGTATTGGTGGGGGAC 939
Qy      274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db      940 ATGTTCAATTACTTCTCTCAGTGGCCATTGGAGTCTTCTGTTCCCGGCCCTCTTACT 999
Qy      294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
Db      1000 GCATCTGCCGTACGTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059
Qy      307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
Db      1060 GAGACAAAGGCAGAGGGCT-----ATCGACTCATCATCAGCGTGGTGGCC 1107
Qy      324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
Db      1108 ATGACTTCACTCTGCTTGTCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATC 1167
Qy      344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGlySer 363
Db      1168 AAAACCCAGGCGAGCCAGTCTACGCCCTCTACCTTGTGCGCCCTCTGCTGTCGACC 1227
Qy      364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db      1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272
```

```
RESULT 24
LOCUS      BC012453      2813 bp      mRNA      linear      PRI 20-AUG-2001
DEFINITION Homo sapiens, clone MGC:9549 IMAGE:3857382, mRNA, complete cds.
ACCESSION  BC012453
VERSION    BC012453.1 GI:15214649
KEYWORDS   MGC.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2813)
AUTHORS    Strausberg,R.
TITLE      Direct Submission
JOURNAL    Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK     NIH-MGC project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgabbs@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```


Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
 Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
 Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
 Newton, N., Nguyen, N., Norris, S., Nwaokelenah, O., Okwuonu, G.,
 Olarpunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
 Perez, A., Perez, L., Plannkoc, C., Plopper, F., Poindexter, A.,
 Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,
 Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
 Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
 Rose, M., Rose, R., Ruiz, J., Sanders, W., Savary, G., Scherer, S.,
 Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
 Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E.,
 Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
 Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
 Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanana, D.,
 Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
 Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R.,
 Worden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
 Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
 Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R.,
 Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 151166)

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GXIX
 Center clone name: CH230-477A14

----- Summary Statistics
 Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 115058 bases at least Q40

Consensus quality: 119558 bases at least Q30

Consensus quality: 122439 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 37 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

* 1 1149: contig of 1149 bp in length

* 1150 1249: gap of unknown length

* 1250 2920: contig of 1671 bp in length

* 2921 3020: gap of unknown length

* 3021 4429: contig of 1409 bp in length

* 4430 4529: gap of unknown length

* 4530 5676: contig of 1147 bp in length

* 5677 5776: gap of unknown length

* 5777 7247: contig of 1471 bp in length

* 7248 7347: gap of unknown length

* 7348 8743: contig of 1396 bp in length

* 8744 8843: gap of unknown length

* 8844 10237: contig of 1394 bp in length

* 10238 10337: gap of unknown length

* 10338 11514: contig of 1177 bp in length

* 11515 11614: gap of unknown length

* 11615 12982: contig of 1368 bp in length
 * 12983 13081: gap of unknown length
 * 13082 14431: contig of 1349 bp in length
 * 14432 14531: gap of unknown length
 * 14532 17021: contig of 2490 bp in length
 * 17022 17121: gap of unknown length
 * 17122 19034: contig of 1913 bp in length
 * 19035 19134: gap of unknown length
 * 19135 20517: contig of 1283 bp in length
 * 20518 22781: contig of 2264 bp in length
 * 22782 22881: gap of unknown length
 * 22882 25072: contig of 2191 bp in length
 * 25073 26957: contig of 1785 bp in length
 * 26958 27058: gap of unknown length
 * 27059 29499: contig of 2342 bp in length
 * 29500 33463: contig of 3964 bp in length
 * 33464 35346: gap of unknown length
 * 35347 35446: gap of unknown length
 * 35447 38703: contig of 3257 bp in length
 * 38704 38803: gap of unknown length
 * 38804 42453: contig of 3650 bp in length
 * 42454 45562: contig of 3009 bp in length
 * 45563 49055: contig of 3403 bp in length
 * 49056 49166: gap of unknown length
 * 49167 51876: contig of 2710 bp in length
 * 51877 54628: contig of 2653 bp in length
 * 54629 58389: contig of 3661 bp in length
 * 58390 58489: gap of unknown length
 * 58490 62913: contig of 4424 bp in length
 * 62914 63013: gap of unknown length
 * 63014 68649: contig of 5636 bp in length
 * 68650 68749: gap of unknown length
 * 68750 76294: contig of 7545 bp in length
 * 76295 81146: gap of unknown length
 * 81147 81246: gap of unknown length
 * 81247 86683: contig of 5437 bp in length
 * 86684 91620: contig of 4837 bp in length
 * 91621 91720: gap of unknown length
 * 91721 99377: contig of 7657 bp in length
 * 99378 108730: contig of 9253 bp in length
 * 108731 120614: contig of 11784 bp in length
 * 120615 120714: gap of unknown length
 * 120715 130786: contig of 10072 bp in length
 * 130787 130886: gap of unknown length
 * 130887 151166: contig of 20280 bp in length.

FEATURES

Location/Qualifiers

source

1. .151166

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone="CH230-477A14"

BASE COUNT 37400 a 32808 c 33338 g 36965 t 10655 others

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

1.24e-34

564.50

58.36%

36.28%

26.43%

2

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

151166

115

70

121

12

5

```
US-09-208-629F-3 (1-407) x AC130634 (1-151166)
Qy 95 ValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 95302 GTTGATGAATCTCTGTCATCGCTCTCACCGGGAAGCTGACCCAGCTCTTCTCCCGGTC 95243
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 95242 ATCTACATCATGCTCTTGTGTAATGTTGGCCAGTAATGATGGCCCTCTGGGTCTTC 95183
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 95182 TTCTTCGCAACGAAGAAGACCCCTGCTGTGATTTTACATGGGCCAACCTGGCCCTGGCA 95123
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 95122 GACCTCTCTCTGTCATCTGTTCCCTCCCTGAAGATCTCCACCACTCCATGGCAACGAC 95063
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 95062 TGGACCTATGGGATGGCTCTGAAGGTGCTCATTTGCTTTTCTACGGCAATATGTAC 95003
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 95002 TGCTCCATCTCTTTTCATGACCTGCTCAGCGTCGAGAGGTACTGGGTGATCGTGAACCCC 94943
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
Db 94942 ATGGGACACTCCAGG---AAGAGGGCCAAACATCTGCTGGCGTCTCCCTGGCCATCTGG 94886
Qy 234 ValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 94885 CTCCTGATTTTCTGGTCACCATCTCTGTACGTATGAGGAGACCATCTACATTTCCA 94826
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 94825 GCCTTGAAACATCACCACTGTCACGAGCTGTGCCCGAG---GAGGTCTGTTGGGGGAC 94769
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293
Db 94768 ATGTTCAGTTACTTCTCTCCCTCCCTGGCCATGGAGTCTTTCTGTTCCAGCCCTCTTACT 94709
Qy 294 IlePheCysTyrThrThrLeuIleHisLysLysSer-----Lys 307
Db 94708 CCGTCTGCTACGTGCTCATGATCAAAACGCTCGTTCCTCCGCCATGGAGCACTCG 94649
Qy 308 AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIle 327
Db 94648 GAGAAGAAAGCGGAGGGCTATCCGCTCATCATCATCAGGTGCTGTCTCATCTTCTATC 94589
Qy 328 CysPheAlaProThrAsnIleIleValIleHisAlaAsnTyrTyrTyrHisAsn 347
Db 94588 TGCTTCGCTCCAGCAACGCTGCTGCTGTGTGATTAATTCCTCATCAAAAGCCAGG 94529
Qy 348 ThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 367
Db 94528 CAGAGCCACGCTACGCGCTCTACTGTCGCGCTCTGCTGTCCACCTCAACAGCTGC 94469
Qy 368 LeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsnPro***Ser 387
Db 94468 ATAGACCCCTTTGTCTACTACTTTGTTTCGAAA-----GATTT-CAGGGACACGAGCAG 94416
Qy 388 AlaMetAlaArgProLeu***ArgProArgArgAspIleTrpGluAspIle 404
Db 94415 AAACCGCGCTCTCTGCCGAAGCGTCCGACCGTGAACCGCATGCAGATATC 94365
RESULT 28
AC010621 AC010621 52368 bp DNA linear PRI 26-JAN-2002
DEFINITION Homo sapiens chromosome 5 clone CTD-2034J20, complete sequence.
ACCESSION AC010621
VERSION AC010621.7 GI:18376859
KEYWORDS HTG.
```

Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 52368)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 52368)
DOE Joint Genome Institute.
Direct Submission
Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 52368)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jan 26, 2002 this sequence version replaced gi:13699549.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence.
Estimated Total Number of Errors is 0.2.
NOTE: This insert is not the entire sequence of the clone (entire
sequence is 142.2kb). It is clipped at the overlap with AC025188.
The number of bases overlapped is 12615.

FEATURES
Location/Qualifiers
1..52368
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2034J20"
BASE COUNT 13910 a 11840 c 12216 g 14402 t
ORIGIN

Alignment Scores:
Pred. No.: 4,16e-35 Length: 52368
Score: 564.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.40% Indels: 28
DB: 7

US-09-208-629F-3 (1-407) x AC010621 (1-52368)
Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 43513 ACCCTTGTCTTCTCTTCTTGTACAGGAACCCAGTAGATCTCTAAAGGAAGACCTTATT 43572
Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
Db 43573 GGTAAAGTTTGAATGGGCATCCCACTGCTGCAAGGAGGAGTTACAGTTGAA----- 43623
Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 43624 -----ACAGTCTTTCTGTGGATGAGTTTCTGCATCTGCTCTCACTGGA 43668
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 43669 AAATGACCACTGCTCTCTCTTCCCAATTTGTCACAAATGTTGTTGGTGGGTTTGCCA 43728
Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
Db 43729 AGTAAGCGCATGGCCCTGTGGGTCTTTCTTTTCCGAACCTAAAGGAAGACCCCTGCTGTG 43788
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
Db 43789 ATTATCATGGCCAACTTGGCCTTGGCTGACCTCTCTCTGTCATCTGGTTCCCTTGAAG 43848
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 165 ILEALATYRHHISLEUASNGLYASNASNTRPVALPHEGLYGLUVALMETCYSARGILETHR 184

```
Db 43849 ATTGCCTATCACATACATACGCAACACTGATTTATGGGAAAGCTCTTTGTAAATGCTT 43908
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
Db 43909 ATTGGCTTTTCTATGCAACATGACTGTTCATCTCTTCATGACCTGCTCAGTG 43968
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 43969 CAGAGGTATTGGGTCACTGTAACCCCATGGGCACTCCAGG--AAGAAGGCAAAACATT 44025
Qy 225 SerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuProPheVal 244
Db 44026 GCATGGCATCTCCCTGGCAATATGCTGCTGATCTGTGTCACCATCTCTTGTAT 44085
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db 44086 GTCGTGAAGCAGACCATCTTCAITTCCTGCTGCAACATCAGACCTGTCTCATGATGTTTG 44145
Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 44146 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 44190
Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeu 300
Db 44191 CTGCGATTGGGTCTTTCTGTTCCCGACCTTCTCTACAGCCTCTGCTATGCTGATG 44250
Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTyrLeuGlyTyr 314
Db 44251 ATCAGAATGCTGGATCTTCTGCCATGATGAACTCAGAGAAAGAAAGAGAGGGCC 44310
Qy 315 IleLysAlaValLeuLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 44311 ATCAAACTATTGTCACTGCTGCGCATGTACCTGATCTGCTTCACTCTAGTAACCTT 44370
Qy 335 IleLeuValIleHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 44371 CTGCTTGTGTGCAATATTTCTGATTAAAGAGCCAGGCGCAGACCATGCTATGCCCTG 44430
Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
Db 44431 TACATTGTAGCCCTCTGCTCTCTACCTTTAAAGCTGATCGACCCCTTGTCTATTAC 44490
Qy 375 ValMetSer 377
Db 44491 TTTGTTTCA 44499

RESULT 29
AR012638
LOCUS AR012638 1255 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5763575.
ACCESSION AR012638
VERSION AR012638.1 GI:3970628
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Sundelin, J. and Scarborough, R. M.
TITLE Agonist and antagonist peptides of the C140 receptor
JOURNAL Patent: US 5763575-A 3 09-JUN-1998;
FEATURES Location/Qualifiers
source 1..1255
/organism="unknown"
BASE COUNT 294 a 320 c 260 g 381 t
ORIGIN

Alignment Scores:
Pred. No.: 7.67e-37 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservatve: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 6 Gaps: 7
```

```
US-09-208-629f-3 (1-407) x AR012638 (1-1255)
Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 116 ACCCTTGCTCTCTTCTTGTACAGGAACCAATAGATCTCTAAAGGAAGAACCTTATT 175
Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
Db 176 GGTAAAGTTGATGGCAGCATCCACAGTCCTGGAAGAGGAGTTACAGTTCAA----- 226
Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
Db 227 -----ACAGTCTTTTCTGTGGATGAGTTTCTGTCATCTGTCTCTCACTGGA 271
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheValValGlyValPro 125
Db 272 AAATGACCACTGCTCTCTCCATTGCTACAAATGTTGTTGTGGTGGTGGTGGCA 331
Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuAthrThr---LysSerIleSerLeuVal 144
Db 332 AGTAACGGCATGGCCCTGTGGGTCTTCTTTCCGAACCTAAAGAAGAACGACCCTGCTGTG 391
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 392 ATTACATGCCAAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGCTGTTCCCTTTGAAG 451
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGCTATCATATACATATGCAACAACTGGATTATGGGGAAGCTCTTTGTAATGCTT 511
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 512 ATTGGCTTTTCTATGGCAACATGACTGTCCATTCTCTCATGACCTGCTCAGTG 571
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 572 CAGAGGTATTGGGTCACTGTAACCCCATGGGCACTCCAGG--AAGAAGGCAAAACATT 628
Qy 225 SerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuProPheVal 244
Db 629 GCATGGCATCTCCCTGGCAATATGCTGCTGATTTCTGCTGTCACCATCTCTTGTAT 688
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db 689 GTCGTGAAGCAGACCATCTTCACTTCTGCTGCTGCAACATCAGACCTGTCATGATGTTT 748
Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 749 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 793
Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeu 300
Db 794 CTGGCCATTGGGTCTTCTTCTTCCAGCCTCTCTCACAGCCTCTGCTATGCTGATG 853
Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTyrLeuGlyTyr 314
Db 854 ATCAGAATGCTGGATCTTCTGCCATGATGAAACTCAGAGAAGAAAGGAAGAGGGCC 913
Qy 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 914 ATCAAACTCATTTGCTACTGCTGCGCATGTACTCATCTGCTTCACTCTCTAGTAACCTT 973
Qy 335 IleLeuValIleHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 974 CTGCTTGTGTGCAATATTTTCTGATTAAAGAGCCAGGCGCAGACCATGCTATGCCCTG 1033
Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
Db 1034 TACATTGTAGCCCTCTGCTCTCTACCTTTAAAGCTGATCGACCCCTTGTCTATTAC 1093
Qy 375 ValMetSer 377
Db 375 ValMetSer 377
```


106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Qy |||||
Db |||||
126 SerAsnIleValThrLeuTyrLysSerLeuArgThr--LysSerIleSerLeuVal 144
Qy |||||
Db |||||
145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Qy |||||
Db |||||
165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Qy |||||
Db |||||
185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyLe 204
Qy |||||
Db |||||
205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Qy |||||
Db |||||
225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Qy |||||
Db |||||
245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Qy |||||
Db |||||
265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Qy |||||
Db |||||
281 LeuAlaPhePheGlyPheLeuProPheValIleIlePheCysTyrThrThrLeu 300
Qy |||||
Db |||||
301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Qy |||||
Db |||||
315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Qy |||||
Db |||||
335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Qy |||||
Db |||||
375 ValMetSer 377
Qy |||||
Db |||||
1094 TTTGTTTCA 1102
Db |||||
RESULT 32
187849
LOCUS 1255 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 3 from patent US 5716789..
ACCESSION 187849
VERSION 187849.1 GI:3407789
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1255)
Sundelin,J. and Scarborough,R.M.
TITLE Method to determine ligands, agonist and antagonist of C140

receptor
Patent: US 5716789-A 3 10-FEB-1998;
FEATURES Location/Qualifiers
1..1255
source /organism="unknown"
BASE COUNT 294 a 320 c 260 g 381 t
ORIGIN
Alignment Scores:
Pred. No.: 7,67e-37 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 6 Gaps: 7
US-09-208-629F-3 (1-407) x 187849 (1-1255)
Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db |||||
116 ACCCTTGTCTTCTTCTTGTACAGGAACCAATAGATCTCTTAAGAGGAAGACGCTTATT 175
Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
Db |||||
176 GGTAAAGGTGATGGACATCCACGTCACCTGGAAAGAGGATTACAGTTGAA----- 226
Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
Db |||||
227 -----ACAGTCTTTTGTGGATGAGTTTCTGCATCTGCTCCTCCTCCTGGA 271
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db |||||
272 AAACCTGACCACTGCTCTTCTTCCAAATTGCTACACAAATTGTTGTGGTGGTTGCCA 331
Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr--LysSerIleSerLeuVal 144
Db |||||
332 AGTAAAGCGGATGCGCTTGGGTCTTTTCCGAACTAAGAAAGAACCCCTGCTGTG 391
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db |||||
392 ATTATACATGCCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGGTTCCCTTGAAG 451
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db |||||
452 ATTGCCTATCACATACATGCAACAACTGGATTATATGGGAAGCTCTTTGTAATGTCTT 511
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyLe 204
Db |||||
512 ATTTGGCTTTTCTATGGCAACATGTACTGTCTCAATTTCTTCATGACCTGCCCTCAGTGTG 571
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db |||||
572 CAGAGGTATTTGGTCTCATCTGAACCCCATGGGGCACTCCAGG---AAGAGGCAACATT 628
Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db |||||
629 GCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGCTCACCATTCTTGTAT 688
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db |||||
689 GTCGTGAAGCAACCACTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db |||||
749 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCTCT 793
Qy 281 LeuAlaPhePheGlyPheLeuProPheValIleIlePheCysTyrThrThrLeu 300
Db |||||
794 CTGGCCATTGGGTCTTTCTGTTCCCAAGCTTCTCTCACAGCCTCTCTCTATGTGCTGATG 853
Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db |||||
854 ATCAGAAATGCTGCGATCTTCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913

[illegible]

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1451)
 Bohm S.K., Kong W., Bromme D., Smeekens S.P., Anderson D.C.,
 Connolly A., Kahn M., Neiken N.A., Coughlin S.R., Payan D.G. and
 Bunnett N.W.
 TITLE Molecular cloning, expression and potential functions of the human
 proteinase-activated receptor-2
 JOURNAL Biochem. J. 314 (Pt 3), 1009-1016 (1996)
 MEDLINE 96177879
 PUBMED 8615752
 REFERENCE 2 (bases 1 to 1451)
 Bohm S.
 TITLE Direct Submission
 JOURNAL Submitted (15-AUG-1995) Stephan Bohm, Department of Surgery, School
 of Medicine, University of California at San Francisco, 521
 Parnassus Avenue, San Francisco, CA 94143-0660, USA
 FEATURES
 Location/Qualifiers
 1..1451
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="kidney"
 1..147
 148..1341
 /note="G-protein-coupled receptor"
 /codon_start=1
 /product="proteinase-activated receptor-2"
 /protein_id="AA84781.1"
 /db_xref="GI:1041729"
 /translation="MRSPSAWLLGAILLAALSLSGTTQGNRRSKGRSLGKVDG
 TSHVTKGVTVETVFSVDFPSAVLTKLTVPLPIVTVFVVGPSNGMALWVFLF
 RTKKHPALVYANWALADLSVWFLPKIAYIHANNWIYGEALCNVLNPLVYKQTIPI
 CSILFWCLSVQRYWVWVNPWHSRKKANIAIGISLAIWLLILVITPLVVKQTIPI
 PALNITCHDVLPEQLVGMFNYFLSLAIGVLPFPAFLTASAYVLMIRLSANDE
 NSKKRRALVITVLVLAFLICFTFENLLVHVFLIKSQGSHVATLVIALCLST
 LNSCIDPFVYFVSHDFRDAKNALLRSVRTVKQMVSLTSKKHSKSSSSYSTT
 VKTSY"
 1342..1451
 310 a 389 c 346 g 406 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.02e-37 Length: 1451
 Score: 563.00 Matches: 119
 Percent Similarity: 52.32% Conservative: 73
 Best Local Similarity: 32.43% Mismatches: 137
 Query Match: 26.36% Indels: 8
 DB: 9 Gaps: 8
 US-09-208-629f-3 (1-407) x HSU34038 (1-1451)
 Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 Db 172 CTGCTGGGGGCGCCATCTCTGCTAGCAGCCTCTCTCTGCGAGTGGCCACCATCCCAAGGA 231
 Qy 42 SerAsnSerAlaLeuLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
 Db 232 ACCAATAGATCCTCTAAGGAGAGCGCTATTGGTAAGTTGATGC----- 279
 Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
 Db 280 -----ACATCCACGTC-----ACTGGAAAAGAGATTACAGTT 312
 Qy 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
 Db 313 GAA-----ACAGTCTTTTCTGGGATGAGTTTCTGCATCT 348
 Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheVal 121
 Db 349 GTCTCTACTGGAACCTGACCACCGTCTCTCCCAATTGCTACAAATTGTTGTG 408
 Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140

Db 409 GTGGGTTTCCCAAGTAACGGCATGGCCCTGGTGGTCTTTCTTTCGAACTAAGAAGAAG 468
 Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 Db 469 CACCCTGCTGTGATTACATGGCCATCTGGCTGGCTGACCTCTCTCTCATCTGG 528
 Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 Db 529 TTCCCTTGAAGATTGCCTATCACATACATGCACTGGCACTGGATTTATGGGAAGCTCT 588
 Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
 Db 589 TGTAAATGCTGCTATTGGCTTTTCTATGGCAACATGCTACTGTTCATTCATGACC 648
 Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 Db 649 TGCTCAGTGTGCAGAGGTATTGGGTATCGTGAACCCCATGGGGCAGCTCCAGG---AAG 705
 Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 Db 706 AAGGCAACATTCGCCATTCCTCGGCAATATGCTGCTGATTTCTGCTGGTCAACC 765
 Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
 Db 766 ATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCTGCAACATCAGCAGCTGT 825
 Qy 261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
 Db 826 CATGATGTTTGT-----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 870
 Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
 Db 871 TACTTCTCTCTGCGCCATTCGGGCTTTCTGTTCCAGCCTTCCTCAGAGCTCTGCC 930
 Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
 Db 931 TATGTCCTGATGATCAAGAAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 990
 Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
 Db 991 AGGAAGAGGCCATCAAACTCATTTGCTGCTGCGCATGATGATCTGCTGCTGCTCACT 1050
 Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
 Db 1051 CCTAGTAACCTTCGCTTGTGGTGCAATTTTCTGATTAAAGAGCCAGGCCAGAGCAT 1110
 Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
 Db 1111 GTCTATGCCCTGTACATTGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGATCGACCCC 1170
 Qy 371 PheLeuTyrPheValMetSer 377
 Db 1171 TTGTCTATTACTTTGTTTCA 1191.
 RESULT 35
 AF400075 18351 bp DNA linear PRI 18-JAN-2002
 LOCUS Homo sapiens coagulation factor II (thrombin) receptor-like 1
 DEFINITION (F2RL1) gene, complete cds.
 ACCESSION AF400075
 VERSION AF400075.1 GI:15021772
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 18351)
 Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L.,
 Yi Q. and Nickerson D.A.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-2001) Molecular Biotechnology, University of
 Washington, 1705 NE Pacific, Seattle, WA 98195, USA

COMMENT	To cite this work please use: SeattleSNPs, NHLBI Program for Genomic Applications, UW-PHCRC, Seattle, WA (URL: http://pga.mbt.washington.edu)	CDS	join(2606, .2687,15958, .17069)
	This sequence consists of 2 contigs. The gap between the contigs is represented as a run of N.		/gene="F2RL1"
FEATURES	10775 10875		/codon_start=1
source	10774: contig of 10774 bp in length 10874: gap of unknown length 10875: contig of 7477 bp in length.		/product="coagulation factor II (thrombin) receptor-like 1"
repeat_region	Location/Qualifiers		/protein_id="AAK77914.1"
variation	1. .18351	variation	/db_xref="GI:15021773"
	/organism="Homo sapiens"		/translation="MRSPSAWLLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG
	/db_xref="taxon:9606"		TSHVTGKGVTVETVFSVDEFSASVLTGKLTVTFLPVTIVTVFVGLPSNGMALWVFLF
variation	343_492	variation	TSKHKPAVIMANALADLLSVINFPKIAIHIGNNWIYGEALCNVLIGFYGNMY
	/rpt_family="MIR"		CSILFMTCLSVORYWVINPMDGHSRKNIAIGISLAIWLLILLVLTPIYVVKQTIFI
variation	/rpt_type=dispersed		PALNITTDHVLPEQLVGDMEFYLSAIGVFLPAPLTAAYVLMIRSSAMDE
	467		NSEKKRKAIKLITVLAWYLICTPTSNLLLVVHYFLIKSQQSQSHVYALYIVALCLST
	/frequency="0.09"		LNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQMVSLTSKHSRKSYSYSSSSTT
	/replace="a"		VKTSY"
variation	531		2667
	/frequency="0.01"		/gene="F2RL1"
	/replace="g"		/frequency="0.02"
repeat_region	678		/replace="t"
variation	728_780		2770
	/rpt_family="L2"		/gene="F2RL1"
	/rpt_type=dispersed		/frequency="0.09"
variation	829		/replace="a"
	/frequency="0.03"		3069
	/replace="a"		/gene="F2RL1"
repeat_region	913_1028		/frequency="0.03"
variation	/rpt_family="FLAM_C"		/replace="a"
	/rpt_type=dispersed		3119
	/frequency="0.01"		/gene="F2RL1"
	/replace="t"		/frequency="0.32"
repeat_region	1059_1214		/replace="t"
variation	/rpt_family="L3"		3151
	/rpt_type=dispersed		/gene="F2RL1"
	/rpt_family="AluSg"		/frequency="0.01"
	/rpt_type=dispersed		/replace="g"
variation	1538		3263
	/frequency="0.01"		/gene="F2RL1"
	/replace="g"		/frequency="0.33"
variation	1622		/replace="g"
	/frequency="0.02"		3306
	/replace="t"		/gene="F2RL1"
repeat_region	1623		/frequency="0.02"
variation	/frequency="0.01"		/replace="t"
	/replace="g"		3330
	/frequency="0.01"		/gene="F2RL1"
	/replace="t"		/frequency="0.01"
variation	1795_1926		/replace="c"
	/rpt_family="MIR"		3331
	/rpt_type=dispersed		/gene="F2RL1"
	/frequency="0.01"		/frequency="0.07"
	/replace="g"		/replace="t"
variation	2122		3353
	/frequency="0.01"		/gene="F2RL1"
	/replace="g"		/frequency="0.01"
variation	2254		/replace="a"
	/frequency="0.01"		3417
	/replace="c"		/gene="F2RL1"
variation	2457		/frequency="0.01"
	/frequency="0.10"		/replace="c"
	/replace="g"		4160
gene	2462_17176		/gene="F2RL1"
mrna	join(2462, .2687,15958, .17176)		/frequency="0.01"
	/gene="F2RL1"		/replace=""
	/product="coagulation factor II (thrombin) receptor-like 1"		4500
variation	2561		/gene="F2RL1"
	/gene="F2RL1"		/frequency="0.02"
	/frequency="0.28"		/replace="t"
	/replace="c"		4663
			/gene="F2RL1"

```

/replacement="0.46"
/replacement="t"
4710.4817
/rpt_family="L2"
/rpt_type=dispersed
4715
/gene="F2RL1"
/frequency="0.15"
/replacement="g"
4972
/gene="F2RL1"
/frequency="0.02"
/replacement="c"
4993.5299
/rpt_family="AluSx"
/rpt_type=dispersed
5065
/gene="F2RL1"
/frequency="0.01"
/replacement="g"
5159
/gene="F2RL1"
/frequency="0.01"
/replacement="a"
5295
/gene="F2RL1"
/frequency="0.01"
/replacement="t"
5296
/gene="F2RL1"
/frequency="0.01"
/replacement="a"
5366
/gene="F2RL1"
/frequency="0.17"
/replacement="t"
5643.5940
/rpt_family="AluSx"

```

Alignment Scores:

```

Pred. No.: 1.55e-35 Length: 18351
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 9 Gaps: 7

```

US-09-208-629F-3 (1-407) x AF400075 (1-18351)

```

Qy 49 ThrLeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 15933 ACCCTTGTCTCTCTTCTGTACAGGAACCAATAGATCCTCTAAAGGAAGCCTTATT 15992

Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
Db 15993 GGTAAAGTTGTAGTCACATCCCGTCACTGGAAGAGGTTACAGTTGAA----- 16043

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 16044 -----ACAGCTTTTCTGTGATGAGTTTCTGCATCTGCTCCTCACTGGA 16088

Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheValValGlyValPro 125
Db 16089 AAACCTGACCACGTCTTCTCCCAATGCTACACAATGTTGTTGTTGGTGGTTGCCA 16148

Qy 126 SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
Db 16149 AGTAACGCATCGCCCTGTGGGTCTTTCTTCCGAACTAAGAAAGACACCTCGCTGTG 16208

Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 16209 ATTTACATGGCAATCTGGCCCTTGGCTGACCTCTCTCTGTCTGTCTGTTCCCTTGAAG 16268

```

```

Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 16269 ATTGCCTATACATATACATGCAACAACTGATTTATGGGAGAGCTTTGTATGATGCTT 16328

Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 18329 ATTGGCTTTTCTATGGCAACATGATGTTCTTCTCATGACCTGCCTCAGTG 16388

Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 16389 CAGAGGTATTGGTGCATCGTAACCCCTGGGCACTCCAGG--AAGAAGGCAACATT 16445

Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 16446 GCCATTGGCATCTCCCTGGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16505

Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 16506 GTCTGTAAGCAGACCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16565

Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 16566 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCTCT 16610

Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeu 300
Db 16611 CTGGCCATTGGGCTCTTCTGTTCCAGCCTTCTCTCACAGCCTCTGCTGCTGCTGCTG 16670

Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 16671 ATCAGAACTGCTGGATCTTCTGCCATGGATGAAACTCAGAGAGAAAGAGAGAGGCC 16730

Qy 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 16731 ATCAAACTCATGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16790

Qy 335 IleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 16791 CTGCTTGTGTGCATTATTTCTGATTAAAGAGCAGGCGCAGAGCCATGCTATGCCCTG 16850

Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuPhe 374
Db 16851 TACATTGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGATCAGCCCTTTGTCTATTAC 16910

Qy 375 ValMetSer 377
Db 16911 TTGTGTTCA 16919

RESULT 36
AC068682/c 184536 bp DNA linear HTG 23-SEP-2000
LOCUS
DEFINITION Homo sapiens clone RP11-206N2, WORKING DRAFT SEQUENCE, 37 unordered
pieces
AC068682
VERSION AC068682.3 GI:10280868
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 184536)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, W., Ferreira, P., Fitzhugh, W., Gage, D.,
Gallagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,

```

Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Laroque, K., Lanazares, R., Landers, T., Lehocsky, J., Levine, R., Lieu, C., Liu, G., Locke, R., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teafaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 23, 2000 this sequence version replaced gi:8247861.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genom Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center Project name: L5802

Center Clone name: 206_N_2

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly: Phrap; version 0.960731

Consensus quality: 167659 bases at least Q40

Consensus quality: 175736 bases at least Q30

Consensus quality: 178914 bases at least Q20

Insert size: 154000; agarose-fp

Quality coverage: 5.0 in Q20 bases; agarose-fp

Quality coverage: 4.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1154: contig of 1154 bp in length
 * 1155 1254: gap of 100 bp
 * 1255 2766: contig of 1512 bp in length
 * 2767 2866: gap of 100 bp
 * 2867 4073: contig of 1207 bp in length
 * 4074 4173: gap of 100 bp
 * 4174 5191: contig of 1018 bp in length
 * 5192 5291: gap of 100 bp
 * 5292 7387: contig of 2096 bp in length
 * 7388 7487: gap of 100 bp
 * 7488 9087: contig of 1600 bp in length
 * 9088 9187: gap of 100 bp
 * 9188 10361: contig of 1174 bp in length
 * 10362 10461: gap of 100 bp
 * 10462 12065: contig of 1604 bp in length
 * 12066 12165: gap of 100 bp
 * 12166 13498: contig of 1333 bp in length
 * 13499 13598: gap of 100 bp
 * 13599 15085: contig of 1487 bp in length
 * 15086 15185: gap of 100 bp
 * 15186 17249: contig of 2064 bp in length
 * 17250 17349: gap of 100 bp
 * 17350 19243: contig of 1894 bp in length
 * 19244 19343: gap of 100 bp
 * 19344 21127: contig of 1784 bp in length

* 21128 21227: gap of 100 bp
 * 21228 23345: contig of 2118 bp in length
 * 23346 23445: gap of 100 bp
 * 23446 25965: contig of 2520 bp in length
 * 25966 26065: gap of 100 bp
 * 26066 28195: contig of 2130 bp in length
 * 28196 28295: gap of 100 bp
 * 28296 32216: contig of 3921 bp in length
 * 32217 32316: gap of 100 bp
 * 32317 36954: contig of 4638 bp in length
 * 36955 37054: gap of 100 bp
 * 37055 40348: contig of 3294 bp in length
 * 40349 40448: gap of 100 bp
 * 40449 43556: contig of 3108 bp in length
 * 43557 43656: gap of 100 bp
 * 43557 47926: contig of 4270 bp in length
 * 47927 48026: gap of 100 bp
 * 48027 54079: contig of 6053 bp in length
 * 54080 54179: gap of 100 bp
 * 54180 58962: contig of 4783 bp in length
 * 58963 59062: gap of 100 bp
 * 59063 63739: contig of 4677 bp in length
 * 63740 63839: gap of 100 bp
 * 63840 71551: contig of 7712 bp in length
 * 71552 71651: gap of 100 bp
 * 71652 78218: contig of 6567 bp in length
 * 78219 78318: gap of 100 bp
 * 78319 84982: contig of 6664 bp in length
 * 84983 85082: gap of 100 bp
 * 85083 91930: contig of 6848 bp in length
 * 91931 92030: gap of 100 bp
 * 92031 98922: contig of 6892 bp in length
 * 98923 99022: gap of 100 bp
 * 99023 106496: contig of 7474 bp in length
 * 106497 108596: gap of 100 bp
 * 108597 114950: contig of 8354 bp in length
 * 114951 115050: gap of 100 bp
 * 115051 124353: contig of 9303 bp in length
 * 124354 124453: gap of 100 bp
 * 124454 132639: contig of 8186 bp in length
 * 132640 132739: gap of 100 bp
 * 132740 141572: contig of 8833 bp in length
 * 141573 141672: gap of 100 bp
 * 141673 151302: contig of 9630 bp in length
 * 151303 151402: gap of 100 bp
 * 151403 161767: contig of 10365 bp in length
 * 161768 161867: gap of 100 bp
 * 161868 184536: contig of 22669 bp in length.

FEATURES

source
 1..184536
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-206N2"
 /clone_lib="RPC1-11 Human Male BAC"
 misc_feature
 1..1154
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 misc_feature
 1255..2766
 /note="assembly_fragment"
 misc_feature
 2867..4073
 /note="assembly_fragment"
 misc_feature
 4174..5191
 /note="assembly_fragment"
 misc_feature
 5292..7387
 /note="assembly_fragment"
 misc_feature
 7488..9087
 /note="assembly_fragment"
 misc_feature
 9188..10161
 /note="assembly_fragment"
 misc_feature
 10462..12065
 /note="assembly_fragment"
 misc_feature
 12166..13498

QY	205	AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe	224
Db	66331	CAGAGTATGGGTATCGTCAATGAGCCCATGGGACATCCAGG--AAGAAGCGCAACATT	66275
QY	225	SerLeuLeuMetCysGlyLeuValTrpValMetValPheLeuTyrMetLeuProPheVal	244
Db	66274	GCCATTGCATCTCCCTGGCAATATGGCTGTGATCTCTGGTCAACATCCCTTTGTAT	66215
QY	245	IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal	264
Db	66214	GTCTGGGAAGCAGACCATCTTCATTCCTGCCCTGAACATCAGCATGTCATGATGTTTGT	66155
QY	265	AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer	280
Db	66154	-----CCTGACGAGCTCTTGGTGGGACATGTTCAATTACTTCTCTCT	66110
QY	281	LeuAlaPhePheGlyPheLeuLeuProPheValIleIlePheCysTyrThrThrLeu	300
Db	66109	CTGGCCATTTGGGTCTTTCTGTCTCCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATG	66050
QY	301	IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr	314
Db	66049	ATCAGAAATGCTGCGATCTTCTGCCATGGATGAAACTCAGAGAAGAAAGAGAGGGGCC	65990
QY	315	IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle	334
Db	65989	ATCAAACTCAITGTACACTGTCTCGCCATGTACCTGATCTGCTTCACTCTCTAGTAACCTT	65930
QY	335	IleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet	354
Db	65929	CTGCTTGTGGTGCATATATTTCTGATTAAAGAGCAGAGCCAGAGCATGCTATGCCCCGT	65870
QY	355	TyrLeuIleAlaLeuLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe	374
Db	65869	TACATTGTAGCCCTCTGCTCTTACCCTTAACAGCTGCATCGACCCCTTTGTCTATTAC	65810
QY	375	ValMetSer 377	
Db	65809	TTTGTTTCA 65801	
RESULT 37			
AC114962		204485 bp	DNA linear
LOCUS		204485 bp	HTG 14-MAR-2002
DEFINITION		8 unordered pieces.	HTG 14-MAR-2002
ACCESSION		AC114962	HTG 14-MAR-2002
VERSION		AC114962.1	HTG 14-MAR-2002
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.	HTG 14-MAR-2002
SOURCE		Homo sapiens.	HTG 14-MAR-2002
ORGANISM		Homo sapiens	HTG 14-MAR-2002
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	HTG 14-MAR-2002
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	HTG 14-MAR-2002
TITLE		DOE Joint Genome Institute.	HTG 14-MAR-2002
JOURNAL		Sequencing of Human Chromosome 5	HTG 14-MAR-2002
REFERENCE		2 (bases 1 to 204485)	HTG 14-MAR-2002
AUTHORS		DOE Joint Genome Institute.	HTG 14-MAR-2002
TITLE		Direct Submission	HTG 14-MAR-2002
JOURNAL		Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint	HTG 14-MAR-2002
COMMENT		Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	HTG 14-MAR-2002
		-----Genome Center	HTG 14-MAR-2002
		Center: Joint Genome Institute	HTG 14-MAR-2002
		Center Code: JGI	HTG 14-MAR-2002
		Web site: http://www.jgi.doe.gov	HTG 14-MAR-2002
		-----	HTG 14-MAR-2002
		Project Information	HTG 14-MAR-2002
		Center Project Name: 485783	HTG 14-MAR-2002
		Center clone name: RPCI-11_206N2	HTG 14-MAR-2002
		-----	HTG 14-MAR-2002
		Summary Statistics	HTG 14-MAR-2002
		Consensus quality: 195215 bases at least	HTG 14-MAR-2002
		O40	HTG 14-MAR-2002

Consensus quality: 199004 bases at least Q30
 Consensus quality: 200415 bases at least Q20
 Estimated insert size: 195000; agarose-fp estimation
 Quality coverage: 9.78 in Q20 bases; agarose-fp estimation.
 Quality coverage: 8.4 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1058: contig of 1058 bp in length
 1059 1158: gap of unknown length
 1159 2342: contig of 1184 bp in length
 2343 2442: gap of unknown length
 2443 4907: contig of 2465 bp in length
 4908 5007: gap of unknown length
 5008 19888: contig of 14878 bp in length
 19889 19986: gap of unknown length
 19987 35850: contig of 15865 bp in length
 35851 53901: gap of unknown length
 53902 53902: contig of 17852 bp in length
 53903 53902: gap of unknown length
 53903 107788: contig of 53886 bp in length
 107789 107888: gap of unknown length
 107889 204485: contig of 96597 bp in length.

FEATURES

source
 1. 204485
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-206N2"
 /clone.lib="RPC1 human BAC library 11"
 BASE COUNT 56752 a 44581 c 44954 g 57495 t 700 others
 ORIGIN

Alignment Scores:

Pred. No.: 2,31e-34 Length: 204485
 Score: 563.00 Matches: 115
 Percent Similarity: 54.52% Conservative: 72
 Best Local Similarity: 33.53% Mismatches: 128
 Query Match: 26.36% Indels: 28
 DB: 2 Gaps: 7

US-09-208-629F-3 (1-407) x AC114962 (1-204485)

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
 Db 73801 ACCCTGTCTTCCTTCTTGTACAGGACCAATAGACCTCTAAAGGAAGAGCCTTATT 73860
 Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
 Db 73861 GGTAAGGTGTAGGCACATCCACGTCACCTGGAAGAGGTTACAGTTGAA----- 73911
 Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
 Db 73912 -----ACAGCTCTTCTGTGTGATGAGTTTCTGTCATCTGCTCCACCTGGA 73956
 Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
 Db 73957 AAACGTGACCATGCTTCCTTCCAAATGCTACCAATGTTGTTGGTGGGTTGCCA 74016
 Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
 Db 74017 AGTAACGCGCATGGCCCTTGGTCTTCTTTTCCGAACTAAGAAGACCCCTGCTGTG 74076
 Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
 Db 74077 ATTTACATGGGCAATCTGGCCTTGGCTGACCTCTCTCTGTCATCTGGTTCCTCCCTTGAAG 74136

Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 Db 74137 ATTGCTATCACATACATAGGCAACACTGGATTATGGGAAGCTCTTTGTAATGCTT 74196
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
 Db 74197 ATTGGCTTTTCTATGGCAACATGACTCTTCCATTCTCTCATGACCTGCCTCAGTGTG 74256
 Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 74257 CAGAGGTATTGGTTCATCTGAACCCCACTGGGGCACTCCAGG---AAGAAGGCAACATT 74313
 Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
 Db 74314 GCCATTGGCATCTCCCTGGCAATATGGCTGCTGATCTCTGCTCACCACCTCTTGTAT 74373
 Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 Db 74374 GTCGTGAAGCAGACCATCTTCTTCTGCTCCAGCCTTCTCCACAGCCTCTGCTATGATGTTTG 74433
 Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
 Db 74434 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 74478
 Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 74479 CTGGCCATTGGGGTCTTCTGTTCCAGCCTTCTCCACAGCCTCTGCTATGCTGATG 74538
 Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
 Db 74539 ATCAGAATCTGCGATCTTCTGCGATGATGAAACTCAGAGAAGAAAGAGAGGGCC 74598
 Qy 315 IleLysAlaValLeuLeuLeuIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 74599 ATCAAACTCATGTCATCTGCTGCGCATGACCTGATCTGCTTCACTCTCTAGTAACCT 74658
 Qy 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
 Db 74659 CTGCTTGTGTCATTATTTCTGATTAGAGCCAGGCGCAGCCATGCTATGCCCTG 74718
 Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
 Db 74719 TACATTGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTGTCTATTAC 74778
 Qy 375 ValMetSer 377
 Db 74779 TTTGTTTCA 74787

RESULT 38

HSU36753 1124 bp DNA linear PRI 28-FEB-1996
 LOCUS Human protease-activated receptor 2 gene, exon 2 and partial cds.
 DEFINITION U36753
 ACCESSION U36753.1 GI:1208539
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1124)
 AUTHORS Kahn,M.L. and Coughlin,S.R.
 TITLE Direct Submission
 JOURNAL Submitted (19-SEP-1995) Mark L. Kahn, CVRI/Medicine, UCSF, 3rd and
 Parnassus, San Francisco, CA 94143-0524, USA
 FEATURES
 source
 1. 1124
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q13"
 /notes="Human genomic DNA in P1 bacteriophage library"
 <3..1112
 /codon_start=1
 CDS

/product="protease-activated receptor 2"
 /protein_id="AAA90957.1"
 /db_xref="GI:1208540"
 /translation="NRSSKGRSLIGKVDGTSHTVGTVEFVSDFASVLTKP
 LTTVFLPIVTVTVVGLPSNGMALVFLPRTKKHPAIVYMANLADLLSVIWFPL
 KIAVHIGNNWIGALCNVLIGFFYGNMYSILEMTCLSVQVYVWVPMGHSRKA
 NIAIGSLMALUILLVLPYVVKOTIFIPALNITTTCHDLVPEQLVGMDFNYFSL
 AIGVFLPAPLTAASVLMIRLSRSDAMSEKKRKAIKLIVTVLWMLYICFTPSN
 LLVAVHFLIKSQSHGVYALYIVLCLSLNSCIDPFVYFVSHDFRDHAKNALLCR
 SVRTVKQMVSLSLTKHSKSSSYSSSSTVKTYS"
 3. .1110
 /product="protease-activated receptor 2"
 /number=2 255 a 292 c 238 g 339 t

exon

BASE COUNT 255 a 292 c 238 g 339 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.08e-36 Length: 1124
 Score: 560.50 Matches: 110
 Percent Similarity: 56.05% Conservative: 66
 Best Local Similarity: 35.03% Mismatches: 113
 Query Match: 26.24% Indels: 25
 DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x HSU36753 (1-1124)

Qy 75 ThrGlyAlaThrThrThrThrLeuLysAlaGluCysProGluAspSerLeuHis 94
 Db 63 ACTGGAAAGAGGTACAGTTGAA-----ACAGTCTTTTCT 98
 Qy 95 ValAsnAlaThrThrThrLeuArgSerSerLeuSerThrGlnValProAla 114
 Db 99 GTGGATGAGTTTCTGTCATCTGCTCCTCAGTGAAGAACTGACCACTGCTCTTCCAATT 158
 Qy 115 IleTyrLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
 Db 159 GTCTACAAATGTGTGTGGGGTGTTCCTCCAGTAACGGCATGCGCTGGGTCTTT 218
 Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
 Db 219 CTTTCCGAACTAAGAGAGACCCCTGCTGTGATTTACATGCGCAATCTGGCGTGGCG 278
 Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
 Db 279 GACCTCTCTCTGTCATCTGCTTCCCTTGAAGATGCTTATCAGTATGCTGCAATGCAAC 338
 Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
 Db 339 TGGATTATGGGGAAGCTCTTTGTAATGCTGCTATTGGCTTTTCTATGTAACATGTATC 398
 Qy 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
 Db 399 TGTTCATCTCTTCATGACCTGCTCAGTGTGCGAGAGGTATTGGGTATCGTGAACCC 458
 Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
 Db 459 ATGGGGCACCCTCAGG---AAGAAGGCAACATGTCATGCGCATCTCCCTGGCAATATGG 515
 Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
 Db 516 CTCTGATCTCTTGGTGCACCATCTTTGTATGTCTGTGAAGAGAGACCATCTTCTCTCT 575
 Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
 Db 576 GCCCTGAACATCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
 Qy 274 -----ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
 Db 621 TTGTGGGAGACATGTTCAATTACTTCTCTCTGCGCATGTTGGGTCTTCTGTTCCCA 680
 Qy 290 PheValIleIleLeuPheCysTyrThrThrLeuIleHisLysLysSer----- 306
 Db 681 GCCTTCTCAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 740

Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
 Db 741 GATGAAACTCAGAGAAAGAAAGAGGCGCATCAAACTCATCTGCTGCTGCGCC 800
 Qy 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
 Db 801 ATGTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
 Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
 Db 861 AAAAGCCAGGCGCAGAGCCATGCTATGCGCTGATGCTAGCTGCTGCTGCTGCTGCTGCT 920
 Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
 Db 921 CTTAACAGCTGCATCGACCCCTTGTCTATTACTTGTGTTTCA 962
 RESULT 39
 AC128225/8955 bp DNA linear HTG 31-JUL-2002
 DEFINITION Rattus norvegicus clone CH230-498A19, *** SEQUENCING IN PROGRESS
 *** 44 unordered pieces.
 AC128225 AC128225.1 GI:21908835
 VERSION HTG; HTGS PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 8955)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbash, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, N., Bryant, N.P.,
 Buahy, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Homi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Hollway, C., Hollins, B.,
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loubege, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okuwonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, C., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 TITLE
 JOURNAL
 Unpublished

REFERENCE

2 (bases 1 to 88955)

AUTHORS

Worley, K.C.

TITLE

Direct Submission

JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 88955)

AUTHORS

Worley, K.C.

TITLE

Direct Submission

JOURNAL

Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GZLT

Center clone name: CH230-498A19

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 35763 bases at least Q40

Consensus quality: 40497 bases at least Q30

Consensus quality: 44674 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
 consists of 44 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

* 1 1023: contig of 1023 bp in length
 * 1024 1123: gap of unknown length
 * 1124 2158: contig of 1035 bp in length
 * 2159 2258: gap of unknown length
 * 2259 3396: contig of 1138 bp in length
 * 3397 3496: gap of unknown length
 * 3497 4759: contig of 1263 bp in length
 * 4760 4859: gap of unknown length
 * 4860 6321: contig of 1462 bp in length
 * 6322 7471: gap of unknown length
 * 7472 7571: gap of unknown length
 * 7572 8781: contig of 1210 bp in length
 * 8782 8881: gap of unknown length
 * 8882 10238: contig of 1357 bp in length
 * 10239 10338: gap of unknown length
 * 10339 11514: contig of 1176 bp in length
 * 11515 11614: gap of unknown length
 * 11615 13207: contig of 1593 bp in length
 * 13208 13307: gap of unknown length
 * 13308 14714: contig of 1407 bp in length
 * 14715 14814: gap of unknown length
 * 14815 16170: contig of 1356 bp in length
 * 16171 16270: gap of unknown length
 * 16271 17526: contig of 1256 bp in length
 * 17527 17626: gap of unknown length
 * 17628 19027: contig of 1401 bp in length
 * 19028 19127: gap of unknown length
 * 19128 20258: contig of 1131 bp in length
 * 20259 20358: gap of unknown length
 * 20359 21562: contig of 1204 bp in length
 * 21563 21662: gap of unknown length
 * 21663 23425: contig of 1763 bp in length
 * 23426 23525: gap of unknown length
 * 23526 24837: contig of 1312 bp in length

* 24838 24937: gap of unknown length
 * 24938 25946: contig of 1009 bp in length
 * 25947 26046: gap of unknown length
 * 26047 27583: contig of 1537 bp in length
 * 27584 27683: gap of unknown length
 * 27684 29343: contig of 1660 bp in length
 * 29344 29443: gap of unknown length
 * 29444 30864: contig of 1421 bp in length
 * 30865 30964: gap of unknown length
 * 30965 32341: contig of 1377 bp in length
 * 32342 32441: gap of unknown length
 * 32442 34492: contig of 2051 bp in length
 * 34493 34592: gap of unknown length
 * 34593 37027: contig of 2435 bp in length
 * 37028 37127: gap of unknown length
 * 37128 40023: contig of 2896 bp in length
 * 40024 40123: gap of unknown length
 * 40124 42440: contig of 2317 bp in length
 * 42441 42540: gap of unknown length
 * 42541 45193: contig of 2653 bp in length
 * 45194 45293: gap of unknown length
 * 45294 47982: contig of 2689 bp in length
 * 47983 48082: gap of unknown length
 * 48083 49458: contig of 1376 bp in length
 * 49459 49558: gap of unknown length
 * 49559 51558: contig of 2000 bp in length
 * 51559 51658: gap of unknown length
 * 51659 53031: contig of 1373 bp in length
 * 53032 53131: gap of unknown length
 * 53132 55703: contig of 2572 bp in length
 * 55704 55803: gap of unknown length
 * 55804 57502: contig of 1699 bp in length
 * 57503 57602: gap of unknown length
 * 57603 59938: contig of 2336 bp in length
 * 59939 60038: gap of unknown length
 * 60039 62314: contig of 2276 bp in length
 * 62315 62414: gap of unknown length
 * 62415 65536: contig of 3122 bp in length
 * 65537 65636: gap of unknown length
 * 65637 68901: contig of 3265 bp in length
 * 68902 69001: gap of unknown length
 * 69002 71414: contig of 2413 bp in length
 * 71415 71514: gap of unknown length
 * 71515 74439: contig of 2925 bp in length
 * 74440 74539: gap of unknown length
 * 74540 78489: contig of 3950 bp in length
 * 78490 78589: gap of unknown length
 * 78590 82200: contig of 3610 bp in length
 * 82201 82299: gap of unknown length
 * 82300 85743: contig of 3444 bp in length
 * 85744 85843: gap of unknown length
 * 85844 88955: contig of 3112 bp in length.

FEATURES

Location/Qualifiers

source

1..88955
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone="CH230-498A19"

BASE COUNT 22306 a 20029 c 19951 g 21690 t 4379 others
 ORIGIN

Alignment Scores:

Pred. No.: 1-75e-34 Length: 88955
 Score: 559.50 Matches: 116
 Percent Similarity: 56.88% Conservative: 66
 Best Local Similarity: 36.25% Mismatches: 121
 Query Match: 26.19% Indels: 18
 DB: 2 Gaps: 6

US-09-208-629F-3 (1-407) x AC128225 (1-88955)

Qy 95 ValAsnAlaThrLeGlyTyrLeuArgSerLeuSerThrGlnValleProAla 114

Db 33815 GTTGATGATTCCTCTGCATCCGTCCTCACCGGAAGCTGACCGCTTTCTCCCGTC 33756

Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeu 134
Db 33755 ATCTACATCATGTCTTGTAAATGGTTTGCAGTAATGGTATGGCTCTGGGTCTC 33696
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 33695 TTCTCCGAAGAGAGAGACCCCTGCTGTGTTTACATGGCCAACTGGGCTTGCCA 33636
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 33635 GACCTCTCTCTGTCATCTGGTCCCTCGAAGATCTCTACCACTCTCATGGCAAGAC 33576
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 33575 TGGACCTATGGGATGGCTCTGCAAGGTGCTCATTTGGCTTTTCTACGGCAATATGTAC 33516
Qy 194 CysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 33515 TGCTCCATCTCTTTCATGAACCTGCTCAGGTGCGAGAGTACTGGGTGATCGTGAACCCC 33456
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIle----- 231
Db 33455 ATGGGACAC-----TCAGGAAGAGGGGCAACATCGTGTGGGTCTCCCTG 33408
Qy 232 ---ValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyr 250
Db 33407 GGCATCTGGCTCTGATTTTCTGGTCACCATCTCTGTACGTGATGAGGAGAACATC 33348
Qy 251 HisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerPro 270
Db 33347 TACATCCAGCTTGAACATCACCATCTGACAGCTGTGCTCCGAG---GAGGTCTG 33291
Qy 271 SerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPhe 290
Db 33290 GTGGGGACATGTTACGTTACTTCTCTCTGCTGCGCATTTGGAGTCTTCTGTTCCAGCC 33231
Qy 291 ValIleIleIlePheCysTyrThrThrLeuIleHisLysLysLysSer----- 306
Db 33230 CTCCTTACTGCTGCTGCTGCTCATGATCAAAACGCTCGCTTCTCCGCCATGGAC 33171
Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIle 324
Db 33170 GAGCACTCGGAGAGAAAGCGGAGGCTATCCGCTCATCATCAGGTGCTGTCCATG 33111
Qy 325 PheThrIleCysPheAlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyr 344
Db 33110 TACTTCATCTGCTTCTGCTCCAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 33051
Qy 345 TyrHisThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeu 364
Db 33050 AGCCAGAGGAGAGCCAGCTACGCTCTACCTGCTGCGCTCTGCTGCTGCTGCTGCTGCT 32991
Qy 365 AsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsn 384
Db 32990 AACAGTGCATAGACCCCTTGTCTACTACTTGTGTTTCGAAA-----GATTT-CAGGGA 32938
Qy 385 Pro***SerAlaMetAlaArgProLeu***ArgProArgAspIleTrpGluAspIle 404
Db 32937 CCAGGCCAGAAACCGCTCTCTCCGAGGCTCGGACCGCTGAACGCGATGCGATATC 32878

RESULT 40

AR012640
LOCUS AR012640 1414 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 62 from patent US 5763575.
ACCESSION AR012640
VERSION AR012640.1 GI:3970630
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1414)
Sundelin,J. and Scarborough,R.M.

TITLE Agonist and antagonist peptides of the C140 receptor

JOURNAL Patent: US 5763575-A 62 09-JUN-1998;

FEATURES

Location/Qualifiers

source 1..1414

BASE COUNT 335 a 361 c 309 g 409 t

ORIGIN

Alignment Scores:
Pred. No.: 8,25e-36 Length: 1414
Score: 551.00 Matches: 118
Percent Similarity: 51.77% Conservative: 72
Best Local Similarity: 32.15% Mismatches: 139
Query Match: 25.80% Indels: 38
DB: 6 Gaps: 8

US-09-208-629F-3 (1-407) x AR012640 (1-1414)

Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
Db 74 CTGCTGGGGGCCGCCATCTCTCTCTAGCAGCCTCTCTCTCTGAGTGGCACCACCAAGGA 133
Qy 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
Db 134 ACCAATAGATCTCTCTAAAGGAAGACCTTATTGGTAAGTTTCTGATGCTC----- 181
Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrIle 81
Db 182 -----ACATCCACGTC-----ACTGGAAAGAGATTACAGTT 214
Qy 82 LysAlaGluCysProGluAspSerIleSerLeuHisValAsnAsnAlaThrIleGly 101
Db 215 GAA-----ACAGTCTTTCTGTGATGATGAGTTTCTGCATCT 250
Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
Db 251 GTCCCTCGCTGGAAACACGACCACTGCTTCCTTCCAAATGCTACACAATTGTTGTTCG 310
Qy 122 ValGlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSer 140
Db 311 GTGGGTTTGCAGATTAACGGCATGGCCTATGGGTCTTCTTTCGAACTAAGAAGAAG 370
Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 371 CACCCTGCTGATTTACATGGCAATCTGGCCTTGGCTGACCTCTCTCTCTCATCTGG 430
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 431 TTCCCTTGAAGATTGCTTATCATACATACATGCAACCACTGGATTTATGGGAAGCTCTT 490
Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
Db 491 TGTAAATGCTTATTTGGCTTTTCTATCGCAACATGATGCTTCCATCTCTCTCATGACC 550
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheTyrThrGlnLysLeuPro 220
Db 551 TGCTCAGTGTGAGAGGATTTGGGTCTGCTGCAACCCCATGGGCACTCCAGG---AAG 607
Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 608 AAGGCAACCAATTCCTTCCATCTCCCTGGCAATATGGCTGTGCTGCTGCTGCTGCTGCT 667
Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
Db 668 ATCCCTTTGTATGCTGTGAGAGACACATCTTCTTCTGCTGCTGCTGCTGCTGCTGCT 727
Qy 261 HisAspValValAspAlaCysGluSerProSerPhe-----ArgPheTyr 276
Db 728 CATGATGTTTG-----CCTGACAGCTCTTGGTGGGAGACATGTTTCAAT 772
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCys 296
Db 773 TACTTCTCTCTCTGCGCCATTTGGGTCTTCTGTTCCAGCCTTCTCTCAGACCTCTGCTGCC 832

```

Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
Db 833 TATGTGCTGATGATCAGAAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAGAAA 892
Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 893 AGGAAGAGGGCCATCAAACTCATGTCTCTGGGCATGTACCTGATCTGCTTCACT 952
Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
Db 953 CCTAGTAACCTTCTGCTGTGGTGCATTATTCTGATTAAAGAGCCAGGCCAGAGCCAT 1012
Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
Db 1013 GTCTATGCCCTGTACATTGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCC 1072
Qy 371 PheLeuTyrPheValMetSer 377
Db 1073 TTGCTCTATTACTTTGTTCA 1093

```

Search completed: June 29, 2003, 09:40:41
Job time : 2691.22 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on:      June 24, 2003, 11:58:31 ; Search time 51.4368 Seconds
              (without alignments)
              1630.373 Million cell updates/sec

Title:       US-09-208-629F-3
Perfect score: 2136
Sequence:    1 TLYTXQHPVAGSQDIKMKIL.....AMARPLXRRRDIWEDIHAW 407

Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5

Searched:    671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	580.5	27.2	399	11	Q8R311	Q8R311 mus musculus
2	429	20.1	374	13	O57466	O57466 meleagris g
3	418	19.6	357	13	Q9DE05	Q9DE05 raja erinac
4	388.5	18.2	361	13	Q90X57	Q90X57 xenopus lae
5	370.5	17.3	309	11	Q8R528	Q8R528 mus musculus
6	369.5	17.3	337	4	Q96P68	Q96P68 homo sapien
7	344.5	16.1	298	11	Q8VE54	Q8VE54 mus musculus
8	342	16.0	377	13	Q8HU14	Q8HU14 brachydanio
9	338	15.8	337	4	O75819	O75819 homo sapien
10	336	15.7	330	4	Q9BXA5	Q9BXA5 homo sapien
11	336	15.7	334	4	Q8TD08	Q8TD08 homo sapien
12	336	15.7	359	11	Q99MT7	Q99MT7 mus musculus
13	335	15.7	359	11	Q9EPD3	Q9EPD3 cavia porce
14	333.5	15.6	358	4	Q98Y21	Q98Y21 homo sapien
15	333	15.6	358	4	Q96JZ8	Q96JZ8 homo sapien
16	330	15.4	351	12	Q9DHV5	Q9DHV5 yaba-like d

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1

Q8R311		PRELIMINARY;	PRT;	399 AA.	
AC	Q8R311				
AD	Q8R311				
DT	01-JUN-2002	(TrEMBLrel. 21, Created)			
DT	01-JUN-2002	(TrEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)			
DE	Coagulation factor II (thrombin) receptor-like 1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
FN	[1]				
RN	SEQUENCE FROM N.A.				
RA	Strausberg R.				
RL	Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; BC025432; AAH25432.1; -.				
KW	Receptor.				
SQ	SEQUENCE 399 AA; 44679 MW; 807C79464AB9B3EF CRC64;				
	Query Match	27.2%; Score 580.5; DB 11; Length 399;			
	Best Local Similarity	37.6%; Pred. No. 1.5e-41;			
	Matches 111; Conservative	64; Mismatches 103; Indels 17; Gaps 5;			
Qy	95 VNNATICGLRSSISTQTVIPAIYILFLFYGVGPSNIIVTLKWLISRT-KSISLVIFHTNLAI	153			
Dd	: :				
Dd	63 IDEFSASILGKTITVLPVVYIIVFVLGPLSPNGMALWIPLFRTPKKHPAVIYMANLALA	122			
Qy	154 DLLFCVTLPKFIAYHLGNMNVGEVMCRITTVVFYGNMYCAILLTCMGINRYLATAP	213			
Dd	: :				
Dd	123 DLLSVWFPLAIAYHLGNMNVGEALCKVLIGFFYGNMYCSILFMFTCLSVQRYYVVI	182			
Qy	214 FTYOKLPRKSFSLLMCQIVMMVFLYMLPFVLKQEYHLVHSEITTHCDVDACESPS	273			
Dd	: ~:				
Dd	183 MGHPR-KKANIAVGSLAIIWLFLVITPIPLYVMKQTIYPALNITTCHDVLPE-EVL	240			
Qy	274 RFYFVSLSAFGFLIPVIIIFCVTTLI-----HKLKSDRIWLGVIKAVLLILV	323			
Dd	: ~:				
Dd	241 MFNYFLSLAICVFLFPAILTATSAVLMIKTLRSAMDHESEKKRQA----IRLIITVLA	296			

```

QY 324 IFTICAPTNIILVIHANYHYHTDSLYFMYLIALCLGSLNSCLDPLFYVMSK 378
DB 297 MYFICAPSNNLLVHFLIKTORQSHVYALYLVLCSTLNSCIDPFYFYVSK 351

RESULT 2
ID 057466 PRELIMINARY; PRT; 374 AA.
AC 057466;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE G protein coupled P2Y nucleotide receptor.
GN TP2Y.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=98086419; PubMed=9415702;
RA Boyer J.L., Waldo G.L., Harden T.K.;
RT "Molecular cloning and expression of an avian G protein-coupled P2Y
receptor.";
RL Mol. Pharmacol. 52:928-934(1997).
DR EMBL; AF031897; AAC60339.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 374 AA; 42594 MW; 849C465722BDD02B CRC64;

Query Match 20.1%; Score 429; DB 13; Length 374;
Best Local Similarity 31.5%; Pred. No. 1.2e-28;
Matches 102; Conservative 56; Mismatches 118; Indels 48; Gaps 10;

QY 74 WTGATTIKAE--CPEDSISTLHVNATIGYLRSSLSSTQVIPAIIYLLFVGVPSNIVTL 131
DB 20 WLGGNTTAAAEAKC-----VFNEEFKFI-----LLPISYGVFVVGGLPLNSWAM 63

QY 132 WKLSRTKS-ISLVIHFTNLAIADLLFCVTLTPFKIAHYHLNGNNVGEVWCRITTVFVG 190
DB 64 WIFVSRMPNMTATTYMFNLAIADLLFVSLPLTVVYADRNWPFQKVFCKIVRELFYA 123

QY 191 NMVCAIILTCMGINRYLATAHPTFYQKLPKRFSLLMCGIWMVFLYMLPVLKQEV 250
DB 124 NLVSSILFUTCSVHRVMGCHPISLKWVTKHARLICVGVWLVVTCILPNIIFVTS 183

QY 251 HLHVSEITTCDDVDVDAESPSSFRFY--YFVSLAFPGFLIPFVIIIFCYTTLIHKL---- 304
DB 184 SKDNS--TLCHDTH---KPEEDHVVHYSSIMALLFGIPFLVIVVCYCLMAKRLCKRS 237

QY 305 -----KSKDRILGYIKAVLLIIVFTICFAPTNIILVIHANYHY----HNTDSL 351
DB 238 FPSPSRPVPSYKKR----SIKMIIVLTVAICFVPHITRTLYTSRYFQADQCLNII 293

QY 352 YFMYLIALCLGSLNSCLDPLFYV 375
DB 294 NFKYKTRPLASINSLDPLIYFM 317

RESULT 3
QYDE05 PRELIMINARY; PRT; 357 AA.
AC QYDE05;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

```

DE P2Y receptor.
OS Raja erinacea (Little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Hymnosquala; Pristiorajae; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7782;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20459151; PubMed=10900200;
RA Dranoff J.A., O'Neill A.F., Franco A.M., Cai S.Y., Connolly G.C.,
RA Ballatori N., Boyer J.L., Nathanson M.H.;
RT "A primitive ATP receptor from the little skate Raja erinacea.";
RL J. Biol. Chem. 275:30701-30706(2000).
DR EMBL; AF242850; AAG42684.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 357 AA; 41239 MW; 14604BE15DCBD41 CRC64;

Query Match 19.6%; Score 418; DB 13; Length 357;
Best Local Similarity 29.4%; Pred. No. 1e-27;
Matches 96; Conservative 62; Mismatches 131; Indels 38; Gaps 9;

QY 62 TFEF--PLSDIEGTGATTIKAECPDSISTLHVNATIGYLRSSLSSTQVIPAIIYLL 119
DB 8 TMEEVFANLSLTGLNKTCS-----LNKGFOFYVLPIMYIV 45

QY 120 FVGVPSNIVTLWKLSRTKS-ISLVIHFTNLAIADLLFCVTLTPFKIAHYHLNGNNVGE 178
DB 46 FVTGFIGNSVLWVFIHFMWSSITTYMFNLVADLFVSLPLIIFYFNKTDWIFGE 105

QY 179 VMCRITTVFVGNMYCAIILTCMGINRYLATAHPTFYQKLPKRFSLLMCGIWMVFL 238
DB 106 LLCKLXRIFHVNLYGSILFTICISVHRVTVGVHPMKSLGRLLKKSATIVCVITVMA 165

QY 239 YMLPVILKQEVHLVHSEITTCDDVDVDAESPSSFRFYFVSLAFPGFLIPFVIIIFCYT 298
DB 166 GISPLIFYSRT-GLRRNKTNCYDTTSK-ELLETY-FIYSMSTTFGFCIPATILVCYG 222

QY 299 TLIHKLKSKDR---IWLGVIKAVLLIIVFTICFAPTNIILVIH-HANYHYHTDS---- 350
DB 223 FIVKALISNDMKTPTURGKSVRLIIVLAVFAISYLPFHVMKNLHQSLRYOGLDTCWN 282

QY 351 --LYFMYLIALCLGSLNSCLDPLFYV 375
DB 283 RVVATYQVTRGLASINSLDPLIYFL 309

RESULT 4
QYDE05 PRELIMINARY; PRT; 361 AA.
AC QYDE05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE P2Y1 nucleotide receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN 1;
RP SEQUENCE FROM N.A.
RA Cheng A.W., Tsim K.W.;
RT "Cloning of Xenopus P2Y1 Receptor.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432354; AAL27614.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.

```


DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to G protein-coupled receptor 23 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019743; AAH19743.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 298 AA; 33708 MW; A09209FBABCE65F2 CRC64;

Query Match 16.1%; Score 344.5; DB 11; Length 298;
Best Local Similarity 30.3%; Pred. No. 1.6e-21;
Matches 76; Conservative 52; Mismatches 88; Indels 35; Gaps 7;

QY 145 IFTNLAIADLLFCVTLPPKIAVHLNGNNVFGVWVCRITTVVFGNMYCAILLTCMGI 204
DB 4 IFTNLALSDLLFCVTLPPKIFNFN-RHWPFGDTLCKSGTAFNLINYGSMFLTCISV 62

QY 205 NRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEVHLVHSEITTCDDVV 264
DB 63 DRELAIVYPPRSRTIRTRNSAIVCAGVILVLSGGISASLF-----STTNVNNA 113

QY 265 DACESPSSRFY--YFVSLAFF-----GFLPFIIVFCVTLIHLK-----SKD 308
DB 114 TTCFEGFSRWKTYLSKITFIETVVGFIPLNLNVSCSVLRLTKPKATLSIGITNKK 173

QY 309 RIWLGYIKAVLLIIVTICFAPTNILVIHANYHYHNTDSLY-----FMYLIALCLGS 363
DB 174 KV-----LKMITHVAVFVGVFVNSVLFYALVRSQAINTCLLERFAKIMVPTLCAT 229

QY 364 LNSCLDPFIYF 374
DB 230 LNCDFDPFIY 240

RESULT 8
Q98U14 ID Q98U14 PRELIMINARY; PRT; 377 AA.
AC Q98U14;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Opioid receptor-like protein ZFOR3.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez R.E., Gonzalez-Sarmiento R., Alvar F.;
RT "Cloning and characterization of ZFOR3, a new zebrafish opioid
receptor-like protein.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF285173; AAG60607.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.

KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 377 AA; 42306 MW; A8AAD07B345064E0 CRC64;

Query Match 16.0%; Score 342; DB 13; Length 377;
Best Local Similarity 27.2%; Pred. No. 3.2e-21;
Matches 86; Conservative 67; Mismatches 125; Indels 38; Gaps 8;

QY 86 PEDSISTLHVNATIGYLRSSLSLQ-----VIPAIYILLFVGVPSNIVTW 132
DB 19 PEKCLPNETWQSAVSDIYNSSLNESWTTEQEAAMSLPIITAVYGVFVGLVGNCLVWY 78

QY 133 KLSLETK---SISLVIFHTNLAIADLLFCVTLPPKIAVHLNGNNVFGVWVCRITTVVY 189
DB 79 VIIIRTKMKTATNIYIF--NLRVADALVTTTFQSTDYLL-NSWPFGEVGVCKVFISIDY 135

QY 190 GNMVCAILLITCMGINRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQE 249
DB 136 YNMFTSIFTLTWMSVDYVAVCHPVKALDFRTPMKAKIINILILWLSAAGIPAMVLGST 195

QY 250 YHLVHSEITTCDDVVDACESPSSRFY--YFVSLAFFGLFPIIVFCVTLIHLKLS 306
DB 196 Q--TNGGTEC-----ALQFPDPYVYVDTLMKICVFIQFVAPLIIITVCYILMVLRLS 248

QY 307 -----KDRIMLGYIKAVLLIIVTICFAPTNILVIHANYHYHNTDSLYPMYLI 357
DB 249 VRLLSGSRKDRNLRITRLVAVVAVFVWCWTPHIFILVKALSHGVPEITTSVMAAYFF 308

QY 358 ALCLGSLNSCLDPPELY 373
DB 309 CVALGYTNSLNPILY 324

RESULT 9
O75819 ID O75819 PRELIMINARY; PRT; 337 AA.
AC O75819;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE T cell death associated protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98317948; PubMed=9655242;
RA Kyaw H., Zeng Z., Su K., Fan P., Shell B.K., Carter K.C., Li Y.;
RT "Cloning, characterization, and mapping of human homolog of mouse T-
cell death-associated gene.";
RL DNA Cell Biol. 17:493-500(1998).
DR EMBL; U95218; AAC31794.1; -
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01649; PSYCHOSINER.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 337 AA; 35339 MW; 05D4739457B079A2 CRC64;

Query Match 15.8%; Score 338; DB 4; Length 337;
Best Local Similarity 28.4%; Pred. No. 6.3e-21;
Matches 94; Conservative 63; Mismatches 118; Indels 56; Gaps 12;

QY 97 NATIGVLRSSLSLQVIPAIYILLFVGVPSNIVTWLKLRLTKSIS-LVIFHTNLAIADL 155
DB 2 NSTCIEEQHDLHLYLFPVIYIFVIIVSIPANIGSLCVSFLOPKKSELGYILFSLSLDL 61

QY 156 LFCVTLPPKIAVHLNGNNVFGVWVCRITTVVFGNMYCAILLITCMGINRYLATAHPPT 215
DB 62 LYALTPLPLWIDYTNKNDWNTFSPALCKGSAFLMYMKFYSSSTAFLTCIAVDYLVVYPLK 121


```
QY 216 YOKLPKRSFSLMCGIWMVFLYMLPFVILKQEXHLVH-----SEITTCDDVVDACS 269
Db 122 FFPLRTRIALMVSLSIWLEITFNA--VLMWEDTWVEYCDAEKSNFTLCYD-----KY 174
QY 270 PSSFRYYFVSLAFF---GFLIPFVILIFC-----YTLIH---KLKSKDRILWLYIKA 317
Db 175 PLE---KQOINLNFCTGAYALPILICNKKVYQAVRHNNKATENKCKRI---IKL 227
QY 318 VLLILVIFTCFAPTNIIIVI-----HANYYYHNTDS--LYFMYLIALCLGSLNSCLDP 370
Db 228 LVSITVFLCFFPHVMLLIRLCILEHAVNFDHNSGKRTYMYRITVALTSLNCVADP 287
QY 371 FLYFVMSKVDQNLNPKSAMARPLXRRDRDI 401
Db 288 ILYCFVTET-----GRYDWM 302

RESULT 10
Q9BXAS PRELIMINARY; PRT; 330 AA.
AC Q9BXAS;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE G-protein coupled receptor 91.
GN GPR91.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Zhang W., Li N., Wan T., Cao X.;
RT "Human P2Y purinoceptor 1.",
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247785; AAL95690.1; -.
SQ SEQUENCE 334 AA; 38697 MW; 33146E1AD87F0E81 CRC64;

Query Match 15.7%; Score 336; DB 4; Length 330;
Best Local Similarity 30.0%; Pred. No. 9.1e-21;
Matches 86; Conservative 61; Mismatches 118; Indels 22; Gaps 8;

QY 105 SSLSTQVIPAIIYLLFVGVGSPSNIVTLWK--LSLRT-KSISLVIFHTNLAIADLLFCVTL 161
Db 19 AALEKYLSIFYGIEFVGVGLGNTIVVGYIFSLKNWSSNIYLF--NLSVSDLAFLCTL 76
QY 162 PKIAYHLNNGNWFGEVWCRTITVVFYGNMYCAIILTCMGINRYLATAHFTYQKLPK 221
Db 77 PMLIRSYANG-NWYIGDVLICSNRYVLHANLYTSILFTFISIDRYLIKIYFPREHLLOK 135
QY 222 RSFSLMCGIWMVFLYMLPFVILKQEXHLVHSEITTCDDVVDACESPSSFRYYFVSL 281
Db 136 KEFALISLAIWLVLTLELP--ILPLINPVITDNGTTCNDP--ASSGDPNLYISMCL 191
QY 282 AFGFLIPFVIIICFTYTLIHKLKSDRIWLG-----IKAVLLILVIFTICFAPTNII 335
Db 192 TLGLFLIPLFVCMCFYKIALFLKQNRQVATAPLEKPLNLVIMAVIFSULFTFYHVM 251
QY 336 LVIIHAN-----YYHNTDSLIFYMILIALCLGSLNSCLDPFLYFVM 376
Db 252 RNVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYELL 298

RESULT 12
Q99MT7 PRELIMINARY; PRT; 359 AA.
AC Q99MT7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G-protein coupled receptor GPR87.
GN GPR87.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
the discovery of new g-protein coupled receptors.",
RL J. Mol. Biol. 307:799-813(2001).
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF295366; AAK01866.1; -.
DR MGD; MGI:1934133; Gpr87.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
```

```
QY 216 YOKLPKRSFSLMCGIWMVFLYMLPFVILKQEXHLVH-----SEITTCDDVVDACS 269
Db 122 FFPLRTRIALMVSLSIWLEITFNA--VLMWEDTWVEYCDAEKSNFTLCYD-----KY 174
QY 270 PSSFRYYFVSLAFF---GFLIPFVILIFC-----YTLIH---KLKSKDRILWLYIKA 317
Db 175 PLE---KQOINLNFCTGAYALPILICNKKVYQAVRHNNKATENKCKRI---IKL 227
QY 318 VLLILVIFTCFAPTNIIIVI-----HANYYYHNTDS--LYFMYLIALCLGSLNSCLDP 370
Db 228 LVSITVFLCFFPHVMLLIRLCILEHAVNFDHNSGKRTYMYRITVALTSLNCVADP 287
QY 371 FLYFVMSKVDQNLNPKSAMARPLXRRDRDI 401
Db 288 ILYCFVTET-----GRYDWM 302

RESULT 10
Q9BXAS PRELIMINARY; PRT; 330 AA.
AC Q9BXAS;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE G-protein coupled receptor 91.
GN GPR91.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
the discovery of new g-protein coupled receptors.",
RL J. Mol. Biol. 307:799-813(2001).
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF348078; AAK29080.1; -.
DR HSSP; P34996; 1DDO.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PS00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G PROTEIN RECP F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 330 AA; 38283 MW; 6D3C70E340866BAB CRC64;

Query Match 15.7%; Score 336; DB 4; Length 330;
Best Local Similarity 30.0%; Pred. No. 9.1e-21;
Matches 86; Conservative 61; Mismatches 118; Indels 22; Gaps 8;

QY 105 SSLSTQVIPAIIYLLFVGVGSPSNIVTLWK--LSLRT-KSISLVIFHTNLAIADLLFCVTL 161
Db 15 AALEKYLSIFYGIEFVGVGLGNTIVVGYIFSLKNWSSNIYLF--NLSVSDLAFLCTL 72
QY 162 PKIAYHLNNGNWFGEVWCRTITVVFYGNMYCAIILTCMGINRYLATAHFTYQKLPK 221
Db 73 PMLIRSYANG-NWYIGDVLICSNRYVLHANLYTSILFTFISIDRYLIKIYFPREHLLOK 131
QY 222 RSFSLMCGIWMVFLYMLPFVILKQEXHLVHSEITTCDDVVDACESPSSFRYYFVSL 281
Db 132 KEFALISLAIWLVLTLELP--ILPLINPVITDNGTTCNDP--ASSGDPNLYISMCL 187
QY 282 AFGFLIPFVIIICFTYTLIHKLKSDRIWLG-----IKAVLLILVIFTICFAPTNII 335
Db 188 TLGLFLIPLFVCMCFYKIALFLKQNRQVATAPLEKPLNLVIMAVIFSULFTFYHVM 247
QY 336 LVIIHAN-----YYHNTDSLIFYMILIALCLGSLNSCLDPFLYFVM 376
Db 248 RNVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYELL 294
```



```

QY 62 TPEPPLSDIEGWTGATTIKACPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFV 121
Db 6 TLAKLNNELHGOESHNSGRSDPGKN-TTLH-----NEEDTIVLPVLYLIIFV 54
QY 122 VGVPSNIVTLW-KLSLRKTSISLVIHTNLAIADLLFCVTLPPKIAHYLNGNNWVGEVM 180
Db 55 ASILLNGLAVWIFFHNRKNT-SFIFYLKNIIVADLLMTLTFPRIVHDAGFGWPYKFFIL 113
QY 181 CRITTVVFGYMYCAIILTCMGINRYLATAHPTTYQKLPKRSFSLMCGIIVVMVFLYM 240
Db 114 CRYTSVLFVANNMTYSIVFLGLSIDRYLKVKVPGDSRMYSITFTKLVSCVVMIVAVLS 173
QY 241 LPVILKQEHVHSEITTCDDVDAC---ESPSSFRFY-----YFVSLAFFGLIPFVII 293
Db 174 LPNII-----LTNQPT--EDNIHDCSKLSPGVKWHATAVTVVNSCLFVAVL---VIL 222
QY 294 IFCYTTLIHLKSKDRIWLG-----YIKAVLLILVIFTICFAPNTIILVIHANNYYH 346
Db 223 IGCYIAISRYIHKSRQFISQSRKRKHQNSIRVVAVFTCTPLPCHLCRIPFTFSHLDR 282
QY 347 NTDS-----LYFMYLIALCLGSLNSCLDPLFLYFVMSK 378
Db 283 LLDESAQKILYCKEITLFLSACNVCLDPIIYFFMCR 319

RESULT 15
Q96JZ8 PRELIMINARY; PRT; 358 AA.
AC Q96JZ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA FLJ14878 fis, clone PLACE1003238, weakly similar to probable G
DE protein-coupled receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamanoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027784; BAB5366.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1;
DR PRINTS; PR01569; F2Y12PRNCPT.
DR PRINTS; PR01655; UDPGLUCOSER.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 358 AA; 41462 MW; 7822700C8E3CC7E9 CRC64;

Query Match 15.6%; Score 333.5; DB 4; Length 358;
Best Local Similarity 25.5%; Pred. No. 1.6e-20;
Matches 86; Conservative 76; Mismatches 132; Indels 43; Gaps 11;

QY 62 TPEPPLSDIEGWTGATTIKACPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFV 121
Db 6 TLAKLNNELHGOESHNSGRSDPGKN-TTLH-----NEEDTIVLPVLYLIIFV 54
QY 122 VGVPSNIVTLW-KLSLRKTSISLVIHTNLAIADLLFCVTLPPKIAHYLNGNNWVGEVM 180
Db 55 ASILLNGLAVWIFFHNRKNT-SFIFYLKNIIVADLLMTLTFPRIVHDAGFGWPYKFFIL 113
QY 181 CRITTVVFGYMYCAIILTCMGINRYLATAHPTTYQKLPKRSFSLMCGIIVVMVFLYM 240

```

```

Db 114 CRYTSVLFVANNMTYSIVFLGLSIDRYLKVKVPGDSRMYNITFTKLVSCVVMIVAVLS 173
QY 241 LPVILKQEHVHSEITTCDDVDAC---ESPSSFRFY-----YFVSLAFFGLIPFVII 293
Db 174 LPNII-----LTNQPT--EDNIHDCSKLSPGVKWHATAVTVVNSCLFVAVL---VIL 222
QY 294 IFCYTTLIHLKSKDRIWLG-----YIKAVLLILVIFTICFAPNTIILVIHANNYYH 346
Db 223 IGCYIAISRYIHKSRQFISQSRKRKHQNSIRVVAVFTCTPLPCHLCRIPFTFSHLDR 282
QY 347 NTDS-----LYFMYLIALCLGSLNSCLDPLFLYFVMSK 378
Db 283 LLDESAQKILYCKEITLFLSACNVCLDPIIYFFMCR 319

RESULT 16
Q9DHV5 PRELIMINARY; PRT; 351 AA.
AC Q9DHV5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 7L protein.
GN 7L.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of
DR EMBL; AU293568; CAC21245.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1;
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 351 AA; 41074 MW; F00A643301887004 CRC64;

Query Match 15.4%; Score 330; DB 12; Length 351;
Best Local Similarity 27.8%; Pred. No. 3.2e-20;
Matches 82; Conservative 73; Mismatches 106; Indels 34; Gaps 13;

QY 91 STLHVNNATIGYLRSSLSSTQVIPAIIYLLFVGVPSNIVTLKLSLRKTSISLV--IFHT 148
Db 39 NVIYDNKNTI-----ISFVYCMFLIFGLLNLVIVLTV-CKLSLSTVTDIFLL 86
QY 149 NLAIADLLFCVTLPPKIAHYLNGNNWVGEVMCRITTVVFGYMYCAIILTCMGINRYL 208
Db 87 NLAIADLLFVMSFPFQIHYQL--SQMIFGNFMCKIVSLGYIIGFSGMFFVTVMVYRI 144
QY 209 ATAHPFTYQ-KLPKRSFSLMCGIIVVMVFLMPLPFVILKQHYLHVHSEITTCDDVDAC 267
Db 145 SIHV-ITVSLKIKTVKIGYISLLLIWIISIVLTPLVVVYQVEK--HDQTLICY----AF 197
QY 268 ESPSSFRFYFVS--LAFPGFLIPFVIIIFCFTTIIHLKLS-KDRIMLGYIKAVLLIIV 324
Db 198 YNNKTFIWLRFNFIENIIGMLIPFVILLFCYVKILMQLCKCKNKNKIKAIRLVLIIFV 257
QY 325 FTICFAPNTIILVI--HHANNYYHNTDSL-----YFMYLIALCLGSLNSCLDPLFLY 373
Db 258 NIIFWIPFNVLFLTSMHSLFEGCKTFKITYALYVTEI-ISSSHCCINPLIY 311

```

RESULT 17

Q9PVY7 PRELIMINARY; PRT; 359 AA.
ID Q9PVY7
AC Q9PVY7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Angiotensin receptor.
OS Anguilla anguilla (European freshwater eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla
NCBI_TaxID=7936;
RN [1]
SEQUENCE FROM N.A.
RA Tran Van Chui M., Dolphin C.T., Barker S., Clark A.J., Vinson G.P.;
RT "Molecular cloning and characterization of the cDNA encoding the
RT angiotensin II receptor of european eel Anguilla anguilla";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AJ005132; CAB40835.1; -;
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000130; Zn_WTpeptidase.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS00290; Ig_MHC; UNKNOWN_1.
DR PROSITE; PS00142; ZINC PROTEINASE; UNKNOWN_1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 39997 MW; 053726A8A827C271 CRC64;

Query Match 15.4%; Score 328; DB 13; Length 359;
Best Local Similarity 25.7%; Pred. No. 4.8e-20;
Matches 82; Conservative 60; Mismatches 109; Indels 68; Gaps 9;

QY 93 LHVNNATIGYRSLSTQVPAIYILLFVVGPSN--IVLWKLSLRKTSISLIVFHTNL 150
DB 12 IHTCNTSG--RHSYIYTLIPVYGCNFGVIGVNSMVAVIYCYMKLTVA-NIFVLN 68
QY 151 AIADLFCVTLPKIAVHLNGNNWVGEVNCRTTTFVFGNMYCAILLTCMGINRYLAT 210
DB 69 AVSDLTFLITLPWATFTAMGYNWPFGGFLCKASAGLTIFNLTSIFLTSLIDRYLAI 128
QY 211 AHPFTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEHVHSEITTC---HDVVD 265
DB 129 VHPVRQRQRTVYARTCTVLIAWAFALSLPTALSRDVTINHPNTVCGTLDKHEL-- 186
QY 266 ACESPSRFYFYFVSL----AFGFLIPFVIIIFCYTTLIHLKSKDRIWLG-----YI 315
DB 187 -----SHVLVAIGLMKSVLGFLLPFIIVTCYCLIGALLAARRVQSSRSRGDEV 237
QY 316 KAVLLILVITICFAPTNILVIHANYHYNTDSLYFMYLIAL----- 359
DB 238 QMLAAVVLAFFLCWVPHQI-----PFHFMVLALLKVIENCPTLDIIDTA 281
QY 360 -----CLGSLNSCLDPLFY 373
DB 282 LPFTICIAFNNSCNPILY 300

RESULT 18

Q9ERK9 PRELIMINARY; PRT; 328 AA.
ID Q9ERK9
AC Q9ERK9
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE P2Y6 receptor (Hypothetical 36.7 kDa protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=129/SVEV;
RX MEDLINE=21160052; PubMed=11259526;
RA Lazarowski E.R., Rochelle L.G., O'Neal W.K., Ribeiro C.M.P.,
RA Grubb B.R., Zhang V., Harden T.K., Boucher R.C.;
RT "Cloning and functional characterization of two murine uridine
RT nucleotide receptors reveal a potential target for correcting ion
RT transport deficiency in cystic fibrosis gallbladder";
RL J. Pharmacol. Exp. Ther. 297:43-49(2001).
RN [2]
SEQUENCE FROM N.A.
RP Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF298899; AAG24619.1; -;
DR EMBL; BC027331; AAH27331.1; -;
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
KW Hypothetical protein; Receptor.
SQ SEQUENCE 328 AA; 36721 MW; 00F9DF5ADADF903E CRC64;

Query Match 15.3%; Score 326; DB 11; Length 328;
Best Local Similarity 27.6%; Pred. No. 6.5e-20;
Matches 86; Conservative 55; Mismatches 109; Indels 62; Gaps 12;

QY 104 RSLSTQVPAIYILLFVVGPSNIVTLWKL--SLRTKTSISLIVFHTNLAIADLFCVTL 161
DB 21 REDFRLILLTPYSVLVVGLPLNCVIAQICASRRLTR--AVYTLNLALADLMYACSL 79
QY 162 PKIAVHLNGNNWVGEVNCRTTTFVFGNMYCAILLTCMGINRYLATAPFP-TVOKLP 220
DB 80 PLLIINYARGDHWPGDGLACRFVRFYANLHGSILFLTCISFQYILGICHPLASWHKRG 139
QY 221 KRSFSLMCGIWMVFLYMLPFVILKQEHVHSEITTCDDVDACESPS--SFRFY-Y 277
DB 140 GREAAMVGVVLAATAVTAQCLPTAVFAATG--IQNRITVCYDL-----SPILSTRILPY 192
QY 278 FVSLAFFGLIPFVIIIFCYTTLIHLKSKDRIWLG-----YIKAVLLILVIF 325
DB 193 GMLATVIGFLPFIALLACYCRMARRLCRD---GPAGPVAQERRSKAARMVAVVAVF 248
QY 326 TICCAPTNILVIHANYHYNTDSLYF-----MYLIALCLGSLNSC 367
DB 249 AISFLP-----FHITKAYLAVRSTPGVSCPVLTPFAAYKGRPFASVNSV 295
QY 368 LDPFL-YFVMSK 378
DB 296 LDPILFYFTQOK 307

RESULT 19

Q8QFR5 PRELIMINARY; PRT; 353 AA.
ID Q8QFR5
AC Q8QFR5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CXCR2 like protein.
GN CXCR2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
RN [1]

QY 349 -----DSLFYFMYLIALCLGSLNSCLDPFLY-FVMSKVVDQL 383
DB 259 RGGKVIETCELRSIDVALVYTOAMAFACAINPILYAFIGKFRNQL 306

RESULT 22

O46685 PRELIMINARY; PRT; 361 AA.
AC O46685;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Orphan G protein-coupled receptor BRGR1B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 15.0%; Score 319.5; DB 6; Length 361;

Best Local Similarity 26.7%; Pred. No. 2.6e-19; Indels 37; Gaps 8;

Matches 76; Conservative 65; Mismatches 107; Indels 37; Gaps 8;
RA Ferrer-Martinez A., Felipe A., Casado J., Pastor-Anglada M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

DR EMBL; U88367; AAC05612.1; -;
DR EMBL; U88366; AAC05611.1; -;
DR HSPSP; P02699; I188.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PRINTS; PR01564; GPCR1RECEPTOR.
DR PRINTS; PR01649; PSYCHOSINER.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 361 AA; 40905 MW; 4802E77C9A45DEE1 CRC64;

Query Match 15.0%; Score 319.5; DB 6; Length 361;

Best Local Similarity 26.7%; Pred. No. 2.6e-19; Indels 37; Gaps 8;

Matches 76; Conservative 65; Mismatches 107; Indels 37; Gaps 8;

QY 113 PAIYILFVVGPSNIVTWKLSIRTKSIS-LVIFHTNLAIADLLFCVTLPFKIAHNG 171
DB 25 PVVYVVLVGFANCLSLYGYLQIKARNEGLVLCNLTVADLFYCSLPFWLQVLQH 84

QY 172 NNWVGEVMCRITTVFVGNMYCAILITCMGINRYLATAHPTFYOKLPKRSFSLMCGI 231
DB 85 DHWSHDDLSCQVCGILLYENIYISVGFLLCCISIDRYLAVAHPPFHQFRTLKAAMGVSA 144

QY 232 VMVYVLYMLPFVILKQEHVHSEITTCDDVDAC-----ESPSFRFYFVSLAFF 284
DB 145 IWKELTSTI-----YFLMHEEVVEDADHRVCFEHPYLPORGINYRFL----V 192

QY 285 GFLPPIFVIFCYTTLIHLK-----SKDRIMLGKIAVLLIIVIFTCFAPTNIL 336
DB 193 GFLPFCILLASRGILRAVRSHGQKSRKQIQ-----RLVLTSTVIFLACPLPHVLL 248

QY 337 VIHANYYYHNTD---SLYFMYLIALCLGSLNSCLDPFLYFVMSK 378

DB 249 LVR--SLWESSCDFAKGIFNAYHFSLLTTSFNCVADPVLVCFVSE 291

RESULT 23

O9UNW8 PRELIMINARY; PRT; 380 AA.
AC O9UNW8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor.
GN G2A.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Weng Z., Witte O.N.;
RT "A DNA damage and stress inducible G protein-coupled receptor blocks
cells in G2/M";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083955; AAD47380.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR01563; G2ARECEPTOR.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1; UNKNOWN 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW Receptor.
SQ SEQUENCE 380 AA; 42499 MW; 6DE63D17275ECD74 CRC64;

Query Match 15.0%; Score 319.5; DB 4; Length 380;

Best Local Similarity 25.5%; Pred. No. 2.7e-19; Indels 41; Gaps 10;

Matches 82; Conservative 67; Mismatches 131; Indels 41; Gaps 10;

QY 73 GWTGATTTIKACPEPDS--ISTLHVNNATIGYLRSSLSSTQVTPAIYILLFVVGVSNIYT 130
DB 9 GYNGNATPVTTTAPWASGLSAKTCTNNVSF-----EESRVLVVVYSVAVCTGLGVANCLT 63

QY 131 LMKLSIRT-KSISLVIFHTNLAIADLLFCVTLPFKIAHNLNNWVFGVMCRITTVY 189
DB 64 AMLALLQVGNLVALYLLCLALCELLYGTGLPLVIYIRNQHRWTLGLLACKVTAYIFP 123

QY 190 GNMVCAILITCMGINRYLATAHPTFYOKLPKRSFSLMCGIWMVFLYMLPFVILKOE 249
DB 124 CMYVSIILFLCCISCDRFVAVVVALESRRRRRTAILISACIFILVGIHVP--VFQTE 181

QY 250 YHLVHSEITTCDDVDACESPSPFRFYFVSLAFFGLIPFVILFICYTTLIHLKSKDR 309
DB 182 -----DKETCFDML---QMSRIAGYVYARFT-VGFAIPLSIIAFTNHRIFRSIKQS-- 229

QY 310 IWLGVYKA-----VLLILVIFTCFAPTNILVIHANYYYHNTD-----SLY 352
DB 230 ---MGLSAAQAKAKVHSAIAVVVIFLVCFAPYHLVLLVKAASFYRGDRNAMCGLEELY 287

QY 353 FMYLIALCLGSLNSCLDPFLY 373

DB 288 TASVVFCLSTVNGVADPIY 308

RESULT 24

O9TTY5 PRELIMINARY; PRT; 342 AA.
AC O9TTY5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
GN PAFR.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.

RA Yang W.S., Diehl J.R., Roubesh W.E.;

RT "Partial Sequence of Bovine Platelet-Activating Factor Receptor

Gene";

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA He B., Triemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,

RA Schmidt P.;

RT "Molecular characterization of bovine platelet-activating factor

RT receptor transcripts and their detection in different tissues of
 RT cattle.",
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF187321; AAF01439.2; -
 DR EMBL; AJ295321; CAC43290.1; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR01559; DUFFYANTIGEN.
 DR PRINTS; PR0237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 342 AA; 39691 MW; 7C4236205AE937C9 CRC64;

Query Match 14.9%; Score 318; DB 6; Length 342;
 Best Local Similarity 27.5%; Pred. No. 3.2e-19;
 Matches 82; Conservative 50; Mismatches 120; Indels 46; Gaps 8;

QY 111 VIPAIYILLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPKIA 167
 Db 17 LPPIFYISIVFVLGVANSYLVWVFARLYPSKKEIKIFMVLNTWADLLFLVTLPLWIVY 76
 QY 168 HLGNNWVFGVWCRITTVVFGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSL 227
 Db 77 YNQGDIWLPKFLCNLAGCFFINTYCSVAFLAVITYNRFQAVTRPIKTAQATTKRGFL 136
 QY 228 MCGIVVMVFLYMLPFVIL---KQYHLVHSEITTCDDVDVADACESPSSFRFYFVSLAF 283
 Db 137 LSLIIWVSIVGAASYFFVLDSTNREPNTKGSANITCFEHEYKSGIPVLTIIHFLV---- 192
 QY 284 FGLPFPVILFCYTLIIHLKSKD-----RIWLGVIKAVLLILVIFTICFAP 331
 Db 193 FSFFLVFLIILFCNLVIIRTLTQQVQIORNAEVRKRLWM-----VCTVLAVFIICFVP 247
 QY 332 TMIILVIH-----ANYVYHNTD---SLVFWYLIALCLGSLNSCLDPPFLYFVMSK 378
 Db 248 -----HHLVQLPWTLAELGQDTHQAINDAHQVTLCLLSTNCVLDPIIYCFLTK 298

RESULT 25

Q9GK76 PRELIMINARY; PRT; 342 AA.
 AC Q9GK76;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Platelet-activating factor receptor.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang W., Diehl J.R., Piumi F.;
 RT "Partial Sequence of Caprine Platelet-Activating Factor Receptor Gene."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF302764; AAG3982.1; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR01559; DUFFYANTIGEN.
 DR PRINTS; PR0237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 342 AA; 39612 MW; 39D9379AED60F4E CRC64;

Query Match 14.8%; Score 317; DB 6; Length 342;
 Best Local Similarity 27.8%; Pred. No. 4e-19;
 Matches 84; Conservative 46; Mismatches 118; Indels 54; Gaps 9;

QY 111 VIPAIYILLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPKIA 167

Db 17 LPPIFYISIVFVLGVANSYLVWVFARLYPSKKEIKIFMVLNTWADLLFLVTLPLWIVY 76
 QY 168 HLGNNWVFGVWCRITTVVFGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSL 227
 Db 77 YNQGDIWLPKFLCNLAGCFFINTYCSVAFLAVITYNRFQAVTRPIKTAQATTKRGFL 136
 QY 228 MCGIVVMVFLYMLPFVILKQYHLVHSE-----ITTCDDVDVADACESPSSFRFYFV 279
 Db 137 LSLIIWVSIVGAASYFFVLDSTNREPNTKGSANITCFEHEYKSGIPVLTIIHFLV 192
 QY 280 SLAFEGFLPFPVILFCYTLIIHLKSKD-----RIWLGVIKAVLLILVIFTI 327
 Db 193 FSFFLVFLIILFCNLVIIRTLTQQVQIORNAEVRKRLWM-----VCTVLAVFI 243
 QY 328 CFAPNTMIILVIH-----ANYVYHNTD---SLVFWYLIALCLGSLNSCLDPPFLYFV 376
 Db 244 CFVP-----HHLVQLPWTLAELGQDTHQAINDAHQVTLCLLSTNCVLDPIIYCFL 296
 QY 377 SK 378
 Db 297 TK 298

RESULT 26

Q9NOU1 PRELIMINARY; PRT; 359 AA.
 AC Q9NOU1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Angiotensin II type 1 receptor.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ADRENAL CORTEX;
 RA Bird I.M., Millican D.S.;
 RT "Isolation of a Full Length Ovine Angiotensin II Type-1 Receptor (AT1-R) cDNA."
 RL Endocr. Res. 0:0-0(2000).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF254119; AAF66063.1; -
 DR HSP; F34996; IDDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
 SQ SEQUENCE 359 AA; 41062 MW; C07010EDB81110EB CRC64;

Query Match 14.7%; Score 315; DB 6; Length 359;
 Best Local Similarity 27.6%; Pred. No. 6.2e-19;
 Matches 85; Conservative 50; Mismatches 103; Indels 70; Gaps 8;

QY 111 VIPAIYILLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPKIA 166
 Db 30 MIPTLYSIIFVVGFGSLWIVIVFYMKLKTVA---SVFLNLALADLCFLTLPLWAV 86
 QY 167 YHLGNWVFGVWCRITTVVFGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSL 226
 Db 87 YTAMEYRWPFGNYLCKIASGSVFNLYASVFLTCLSDRYLAIIVHPMKSRRLRTMLVAK 146
 QY 227 LMCGIVVMVFLYMLPFVILKQYHLVHSEITTCDDVDVADACESPSSFRFYFV----- 277
 Db 147 VTCIIILWLAGLASLPTIIRNVFFIENTNITVC-----APHYSQNSTLPV 193
 QY 278 --FVSLAFFGLPIIPVILFCYTLIIHLKLS-----KDIRWLGVIKAVLLILVIF 325

Db 194 GLGLTRNKLGLFPFLPFIILTSYTLWTKLKAYEIQNKPRKDDIP---KIILAIVLFP 249
Qy 326 TICFAPTNILVTHHANYHYNTDSLYFMYL-----IALCLGSLNSCLDP 370
Db 250 FFSWPHQIPTF-----DVLQILRDCKIEDIVDTAMPITICLAYFNCLNP 299
Qy 371 FLYFVMSK 378
Db 300 LFYGFGLK 307

RESULT 27
Q8QG04 PRELIMINARY; PRT; 390 AA.
AC Q8QG04;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Type five-like somatostatin receptor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Peter R.E.;
RT "Cloning and expression of a type five-like somatostatin receptor in goldfish."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252879; AAM18805.1; --
KW Receptor.
SQ SEQUENCE 390 AA; 44136 MW; EC12119A4B6CF9A8 CRC64;

Query Match 14.6%; Score 311; DB 13; Length 390;
Best Local Similarity 28.4%; Pred. No. 1.5e-18;
Matches 80; Conservative 64; Mismatches 116; Indels 22; Gaps 7;

Qy 108 STQVIPAIVYLLFVGVPSNIVTLWKLRLTK-SISLVIFHTNLAIADLLFCVTLPPFKIA 166
Db 39 STKILAVILVFWGVGLNSLAIFVFLRYTKMTATNNYILNLAVADELYLGLPFLTA 98
Qy 167 YHLGNWNVFGEVCMCRITTVFVGNMYCAIILTCMGINRYLATAHPFTYQKLPKRSFSL 226
Db 99 HNMVG-YWFGNPLCLILMKWTDISQSTFTCLTVMSIDRYAVVPIRSARWRPRVAK 157
Qy 227 LMGIVVMVFLVYLPVILKQYHLVHSEITFCHDVVDACSPSSFRFYFVSLAFFGF 286
Db 158 VINSWVWALSCLLTLFVII---YCDVQPELNTCN--LSWPEPRDVMSTAFILYTMALGF 211
Qy 287 LIPFVLIICYYTLIHLKLS-----KDRWLGYIKAVLLIIVFTICFAPTNIILV 337
Db 212 FFLPMVICLYLLIVIKVSASARAGLSKRGSEKVKTRMVIIVVFLVCLWLPFFIINI 271
Qy 338 IHHANYHYNT--DSLVFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 272 LNLISLTPENSLMTGIFYLTVI---LTYVNSCANPLLYSFLS 310

Q9XSD4 PRELIMINARY; PRT; 342 AA.
AC Q9XSD4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]

RESULT 28

Q9XSD4 PRELIMINARY; PRT; 342 AA.
AC Q9XSD4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]

RP SEQUENCE FROM N.A.
RA Diehl J.R., Murphy K.E., Roudabush W.E.;
RT "porcine (Sus scrofa) platelet-activating factor receptor DNA";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF124054; AAD28739.2; --
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; UNKNOWN 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1.2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 39435 MW; 830EA7542B528DAC CRC64;

Query Match 14.5%; Score 310; DB 6; Length 342;
Best Local Similarity 27.5%; Pred. No. 1.6e-18;
Matches 81; Conservative 50; Mismatches 124; Indels 40; Gaps 8;

Qy 111 VIPAIYILLFVGVPSNIVTLW---KLSLRTKSISLVIFHTNLAIADLLFCVTLPPFKIA 167
Db 17 LFFIPYSIIFVLGVANSVLMVFPARVYPSKLNKIKIFMLNLTWADLLFLVTLPLWIY 76
Qy 168 HUNGNNWVGEVCMCRITTVFVGNMYCAIILTCMGINRYLATAHPFTYQKLPKRSFSL 227
Db 77 YYHEGNWILPKFLCNLAGCFEINTYCSVFLAVIYTRFQAVTRPIKTAQATTKRGIS 136
Qy 228 MCGIVVMVFLVYLPVILKQYHLVHS---EITTDHVDVDCSPSSFRFYFVSLAF 283
Db 137 LSLIILWAVMAAASVFFVLDSTNIELSKTGAGNLTRCFEHYKSGMPULIHLFV--- 192
Qy 284 FGLPFIPIIIFCYTTLIHLKLSK---RILWGYIKAVLLIIVFTICFAP 331
Db 193 FSEFLVFLVILFVILVITLTLTQSQVQMRNAEVKRRALWM---VCTVLAVFICFVP 247
Qy 332 TNII-----LVTHANYHYNTDSLYFMYLIALCLGSLNSCLDPFLYFVMSK 378
Db 248 HHIVQLPWTABLPGSQGNFHOAINDA---HQITLCLLSTNCVLDPIIYCFITK 298

RESULT 29
Q9GLN9 PRELIMINARY; PRT; 359 AA.
AC Q9GLN9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Angiotensin II type-1 receptor.
OS AGTRI.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20469400; PubMed=11013071;
RA Dufour C., Casane D., Denton D., Wickings J., Corvol P.,
RA Jeunemaitre X.;
RT "Human chimpanzee DNA sequence variation in the four major genes of the renin angiotensin system."
RL Genomics 69:14-26(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF193445; AAG28410.1; --
DR HSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1.2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 41033 MW; 266133A7A3E911A6 CRC64;

Query Match 14.5%; Score 310; DB 6; Length 359;

[illegible]


```

QY 302 ----HKLKGRWLGVKAVLLILVITFCAPTNIILVI-----HHANYYYHNTDSLY 352
DB 227 LRRPNKKSK-----AVRLFVIMIIFFLWTFPNYLLTILSVFQDFLFYTCCEQGRQLD 280

QY 353 FMYLIALCLGSLNSCLDPFLY 373
DB 281 LAIQVTETMIATYHCCVNPVY 301

RESULT 37
Q9BDS6 PRELIMINARY; ~ PRT; 343 AA.
AC Q9BDS6;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Orphan seven transmembrane receptor STRL33.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21134756; PubMed=11242524;
RA Wade-Evans A.M., Russell J., Jenkins A., Javan C.;
RT "Cloning and sequencing of cynomolgus macaque ccr3, gpr15, and strl33:
RT potential coreceptors for HIV type 1, HIV type 2, and SIV."
RL AIDS Res. Hum. Retroviruses 17:371-375(2001).
DR EMBL; AF291671; AAK25742.1; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; LYMPHOTACTNR.
DR PROSITE; PS00262; G PROTEIN RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G PROTEIN RECEPTOR_F1_2; 1.
KW Receptor; Transmembrane.
SQ SEQUENCE 343 AA; 39472 MW; 0961328F948E7784 CRC64;

Query Match 14.3%; Score 304.5; DB 6; Length 343;
Best Local Similarity 26.9%; Pred. No. 4.6e-18;
Matches 87; Conservative 62; Mismatches 108; Indels 67; Gaps 14;

QY 112 IPAIYILLFVVGPSNIVTLWLKSLRTKSLV-IFHTNLATADLLFCVTLPPKIAHYLN 170
DB 35 LPCMILVFCVGLGNSLVLSIFVHKQLQSLTDFVNLPLADLVFVCTLPFWT--YAG 92

QY 171 GNNVFGVCMCRITTVVFGNMYCAIILTCMGINRYLATAHPF-TYQKLPKR-SFSLM 228
DB 93 IHEWIFQVCKTLLGVYTNFTYMSMLTCTIVDRFVVVKATKAYNQAKEMTWGKVI 152

QY 229 CGIWWNVFLMPLFVLKQEYHL-----VH-SEITTHDVVDACESPSSFRFYFVSL 281
DB 153 CLLIWIISLVSLPQIYIGNVFNLDKLCGYHDEEIST--VVLATQMT----- 198

QY 282 APTGELIPFVIIIFCYTLLHLK-----SKRIWLGVIKAVLLILVITFCAPTNIIL 336
DB 199 --LGFLPLAMIVCVSVIKTLHLHAGGFKHR-----SLKIFLVMAVLLTQTPNNVK 252

QY 337 VIHANYYYHNTDSLVMYLIACLSLNSCLDPFLY-FVMSKVVDQLNPXSAMARPLXR 395
DB 253 LIRSTRWEYAMTSFHYTIIVTETAIYLRACLNPLVYAFVSLK----- 296

QY 396 PRDIW---BDI-----HAW 407
DB 297 -RKNFWKLKVDIGCLPLVGLVSHOW 319

RESULT 38
O42324 PRELIMINARY; PRT; 383 AA.
ID O42324
AC O42324;
```

```

DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Mu-opioid receptor.
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=CNS;
RC MEDLINE=97368346; PubMed=9223341;
RA Darlison M.G., Greten F.R., Harvey R.J., Kreienkamp H.J., Stuenkel T.,
RA Zwiers H., Lederis K., Richter D.;
RT "Opioid receptors from a lower vertebrate (Catostomus commersoni):
RT sequence, pharmacology, coupling to a G-protein-gated inward-
RT rectifying potassium channel (GIRK1), and evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 94:8214-8219(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; Y10904; CA471843.1; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR_F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECEPTOR_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 383 AA; 43232 MW; F58838B57A107305 CRC64;

Query Match 14.2%; Score 304; DB 13; Length 383;
Best Local Similarity 28.9%; Pred. No. 5.7e-18;
Matches 94; Conservative 61; Mismatches 108; Indels 62; Gaps 14;

QY 88 DSISTLHVNNATIGYLRSSLSLSTQ-----VIPATVILLFVVGPSNIVTLWLKSLRT 138
DB 31 NSTSFLNN-----GSSRSDSTDEQDKTPVIAIIITLTSIVCVGLGVNLVYVIIRYT 86

QY 139 K--SISIVIPHTNLAIADLLFCVTLPPKIAVHLGNWVFGVCMCRITTVVFGNMYCA 195
DB 87 KMKATNIYIP--NLALADALATSTLPQSVNYLNG-TWPGDVGVCKIWMSDIYNMFTS 143

QY 196 ILILTCMGINRYLATAHPFTYQ--KLPRKSFSLMCGIWMVFLMPLFVLKQEYHLV 253
DB 144 IFTLTMSIDRYIAVCHPVKALDPRTPRNKIVNVCN--WILSSAIGLPVMVMAS----- 196

QY 254 HSEITTHDVVDACESP--SSFRF-----YFVSLA-----FFGFLIPFVLIICY 297
DB 197 ----TT-----TENQNSPLQVSNFDCITLLFPHPPWYMETLLKTCVFILAFIMPVLIITCY 248

QY 298 TTLHLKLS-----KDRILWLGVIKAVLLILVITFCAPTNIILVIHANYYYHNT 348
DB 249 GLMIURLKSVMLSSGSKEDRNLRITRWLVVVAVFICWTPHIFVII-KALVTIPNS 307

QY 349 DLSYFMYLIACLSLNSCLDPFLY 373
DB 308 LFQVTWHFCIALGYTNSCLNPVLY 332

RESULT 39
Q9JLZ0 PRELIMINARY; PRT; 362 AA.
AC Q9JLZ0;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Orphan chemokine receptor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC TISSUE=MUSCLE;
RA Xie P., Fu A.K.Y., Ip N.Y.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AF118816; AAF34338.1; -.
DR HSSP: P34996; 1DD.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm1.1.
DR PRINTS: PR00237; GPCRHHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 362 AA; 41650 MW; C47D1DB678697CFD CRC64;

Query Match      14.2%; Score 303.5; DB 11; Length 362;
Best Local Similarity 25.8%; Pred. No. 5.9e-18;
Matches 73; Conservative 62; Mismatches 117; Indels 31; Gaps 8;

QY 115 YILLFVVGVPNIVTLW-KLSLRKTSISLVPHNTLAIDLFCVTLFPKIAHYHLNGNN 173
Db 115 YILLFVVGVPNIVTLW-KLSLRKTSISLVPHNTLAIDLFCVTLFPKIAHYHLNGNN 173
QY 50 IYIFIFVIGMIANSVVVWNIQAKTTGYDTHCYILNLAIDLAWVITIPVWVSVLQHNQ 109
Db 50 IYIFIFVIGMIANSVVVWNIQAKTTGYDTHCYILNLAIDLAWVITIPVWVSVLQHNQ 109
QY 174 WVFGEWMCRIITVVFYGNMYCAILITCMGINRYLATAHPFTYQKLPKRSFSLMCGIVM 233
Db 174 WVFGEWMCRIITVVFYGNMYCAILITCMGINRYLATAHPFTYQKLPKRSFSLMCGIVM 233
QY 110 WPMGEUTCKITHLIFSINLFGSIFFLACMSVDRLYSITYFTSTSSYKKQVRRVVCVLVM 169
Db 110 WPMGEUTCKITHLIFSINLFGSIFFLACMSVDRLYSITYFTSTSSYKKQVRRVVCVLVM 169
QY 234 VMVFLYMLPFVILKQEHVHSEITTCDDVDACESPSFRFY-----YFVSL----A 282
Db 234 VMVFLYMLPFVILKQEHVHSEITTCDDVDACESPSFRFY-----YFVSL----A 282
QY 170 LLAF-----FVSLPDYTL--KTVTSASNNETVCRS-----FYPEHSIKEWLIGMELSV 217
Db 170 LLAF-----FVSLPDYTL--KTVTSASNNETVCRS-----FYPEHSIKEWLIGMELSV 217
QY 283 FFGFLIPFVIIIFCYTTLIHKLK-SKDRIWLGVIKAVLLILVIFTICFAPTNIIL----- 336
Db 283 FFGFLIPFVIIIFCYTTLIHKLK-SKDRIWLGVIKAVLLILVIFTICFAPTNIIL----- 336
QY 218 ILGFVAPFTIIAIFYELLARAMSAGDQEKHSRKIIFSVWVFLVCWLPYHFVLLDIF 277
Db 218 ILGFVAPFTIIAIFYELLARAMSAGDQEKHSRKIIFSVWVFLVCWLPYHFVLLDIF 277
QY 337 -VIHANYHYNTDSLYFMVLTALCLGSLNSCLDPFLYFVMSK 378
Db 337 -VIHANYHYNTDSLYFMVLTALCLGSLNSCLDPFLYFVMSK 378
QY 278 SILHYIPTCQENLVLTALHVTQCLSLVHCCVNPVLYSFINR 320
Db 278 SILHYIPTCQENLVLTALHVTQCLSLVHCCVNPVLYSFINR 320

RESULT 40
Q96HI6 PRELIMINARY; PRT; 331 AA.
AC Q96HI6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:17339).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008569; AA08569.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm1.1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 331 AA; 38133 MW; 5B0B755A74FBB36 CRC64;

Query Match      14.2%; Score 303; DB 4; Length 331;
Best Local Similarity 30.2%; Pred. No. 6e-18;
Matches 83; Conservative 55; Mismatches 113; Indels 24; Gaps 11;

QY 116 YILLFVVGVPNIVTLWKLRLKTSISLVPHNTLAIDLFCVTLFPKIAHYHLNGNN 174
Db 116 YILLFVVGVPNIVTLWKLRLKTSISLVPHNTLAIDLFCVTLFPKIAHYHLNGNN 174
QY 29 YSCIFIIGLVNITALWVFCSTTKKRTTIVYMNVALVDLIFIMLPFRMFYAK-DEW 87
Db 29 YSCIFIIGLVNITALWVFCSTTKKRTTIVYMNVALVDLIFIMLPFRMFYAK-DEW 87
QY 175 VFGEVNCRI--TTVFGNMYCAILITCMGINRYLATAHPFTYQKLPKRSFSLMCGIV 232
Db 175 VFGEVNCRI--TTVFGNMYCAILITCMGINRYLATAHPFTYQKLPKRSFSLMCGIV 232

```

```

Db 88 PEGEYFCQILGALTVPFYPSI--ALWLLAFISADRYMAIVQPKYAKELKNTCKAVLACVG 145
QY 233 WMVFLYMLPFVILKQEHVHSEITTCDDVDACESPSFRFYFVSLAFTGFLIPFVI 292
Db 146 WIMTLITTTTPLLILYKDDP-KDSTATCLKISDIYL-KAVNVLNLTTLTFF-FLIPLFI 202
QY 293 IIFCYTTLIHKL-----KSKDRIWLGVIKAVLLILVIFTICFAPTNI---ILVIHHANY 343
Db 203 MIGCYLVIITHNLLHGRTSKLKPKEKESIRIITLLVQVLVCFMPFHCFAFLMLGTGEN 262
QY 344 YHNTDSLYFMVLTALCLGSLNSCLDPFLYFVMSK 378
Db 263 SY-NPWGAFTTFLM-----NLSTCLDVILYIVSK 291

Search completed: June 24, 2003, 12:05:35
Job time : 55.4368 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:59:21 ; Search time 25.9681 Seconds
(without alignments)
1506.722 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXXHPVAGSODIKKIL.....AMARPLXRRRDIWIDIAHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582.5	27.3	420	2 I51667	thrombin receptor
2	581.5	27.2	399	2 I48705	proteinase activat
3	569	26.6	397	2 S66518	proteinase-activat
4	521	24.4	427	2 S17148	alpha-thrombin rec
5	512.5	24.0	432	2 A43448	thrombin receptor
6	509	23.8	425	2 A37912	thrombin receptor
7	431.5	20.2	361	2 B45680	G protein-coupled
8	399	18.7	362	2 S33733	G protein-coupled
9	389.5	18.2	370	2 JC5549	heptahelical P2Y5-
10	383	17.9	308	2 I50241	G protein-coupled
11	378.5	17.7	373	2 A47556	ATP receptor P2u -
12	376.5	17.6	373	2 JC4737	G protein-coupled
13	372	17.4	373	2 JC4162	P2Y receptor - bov
14	362	16.9	365	2 S68679	G protein-coupled
15	357	16.7	380	2 JC2434	kappa opioid recep
16	357	16.7	380	2 A48227	kappa opioid recep
17	356	16.7	344	2 T09508	intrin 17 purinerg
18	355.5	16.6	362	2 JN0694	angiotensin II rec
19	355	16.6	380	2 I38435	angiotensin recept
20	353.5	16.5	380	2 JC2338	kappa opioid recep
21	352	16.5	380	2 S36143	kappa opioid recep
22	350	16.4	380	2 A55259	kappa opioid recep
23	347.5	16.3	375	2 A54946	P-2U nucleotide re
24	333.5	15.6	362	2 B57641	G protein-coupled
25	333	15.6	358	2 I55450	G protein-coupled
26	332	15.5	329	2 I51372	angiotensin II rec
27	327	15.3	359	2 JC2134	angiotensin II rec
28	326	15.3	359	2 JC0621	angiotensin II rec
29	325	15.2	353	2 JC2492	G protein-coupled

30	324.5	15.2	372	2 S26667	G protein-coupled
31	324	15.2	359	2 JC1194	angiotensin II rec
32	321.5	15.1	354	2 I53033	G protein-coupled
33	321.5	15.1	362	2 A57641	G protein-coupled
34	320	15.0	359	2 S44425	angiotensin II rec
35	320	15.0	359	2 A48857	angiotensin II rec
36	320	15.0	359	2 S15403	angiotensin II rec
37	319.5	15.0	361	2 JC5653	G protein-coupled
38	318.5	14.9	365	2 S68208	G protein-coupled
39	318	14.9	358	2 A53752	interleukin-8 rece
40	317.5	14.9	328	2 JC4800	P2Y6 receptor - hu
41	317.5	14.9	341	2 S63666	platelet activatin
42	317.5	14.9	362	2 S68207	G protein-coupled
43	315	14.7	359	2 A42656	angiotensin II rec
44	314	14.7	355	2 A55733	G protein-coupled
45	314	14.7	355	2 A45177	chemokine (C-C) re

ALIGNMENTS

RESULT 1

I51667
Thrombin receptor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51667
R;Gerstzen, R.E.; Chen, J.; Ishii, M.; Ishii, K.; Nanevicz, T.; Turck, C.W.; Vu, T.H.;
Nature 368, 648-651, 1994
A;Title: Thrombin receptor's specificity for agonist peptide is determined by its extra
A;Reference number: I51667; MUID:94195429; PMID:8145852
A;Accession: I51667
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-420 <GER>
A;Cross-references: EMBL:U09632; NID:G495197; PIDN:AAA18498.1; PID:G495198

Query Match 27.3%; Score 582.5; DB 2; Length 420;

Best Local Similarity 32.7%; Pred. No. 1.2e-38;

Matches 124; Conservative 82; Mismatches 138; Indels 35; Gaps 11;

QY	19	ILILVAAGLFLPVTVCQSGINVSNAK-----PTLTIKSFN--GGPQNTFEFFPLSDIE	72
DB	7	LLLLLLLTLLGAMGSLCLAN---SDTQAKGAHNNMTIKTRIFDSESEFEIIPWDEL	63
QY	73	GMTGATTTIKAECPEDSISTL--HVNNTATIGVLRSSLSTQVIPAIYILLVWGVPSN---	127
DB	64	E-SGEGSQQAPVRSRARKPIRNTKEAEQVLSQWLTKFVPSLTVTVFVIGLPLNLLA	122
QY	128	-IVTLWKLRLTKSISLVIFHTNLAIADLLFCVTLFPFKIAYHLNGNNWVGEVYMCRTTV	186
DB	123	IIIFLFPMKVRKPA---VVYMLNLAIADVFVSVLPFKIAYHLNSGNDWLFPGCMCRIVTA	179
QY	187	VFYGNMYCAILLTCMGINRYLATAPFTYQKLPKRPSLLMCGIWWVFLYMLPFVIL	246
DB	180	IFYCNNYCVLLIASISVDRLAVVYPMHSLSWRTMSRAYMACSFILWISIASTIPLLV	239
QY	247	KQEHYHLVHSEITTCDDVDACSPSSFRFYVFLSAFFGLFIPFVLIIFCYTTLHLKLKS	306
DB	240	EQTKIPRDIITTCDDVLD-LKDLKDFIYYIYSSFCLLFFVFPIITTCIYIGIISLSS	298
QY	307	-----KDRILWGLYKIKAVLLILVITTCFAPTNIILVIHANYYYHNTDSLYFYMLIA	358
DB	299	SSIENSCCKTRA---LFLAVVLCVFIICFGTNTVFLTH---YLQEAERFLYFAYILS	351
QY	359	LCGLNSCLDPFLYFVMS	377
DB	352	ACVGSVSCCLPLIYYIAS	370

RESULT 2

I48705
proteinase activated receptor 2 - mouse

C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I48705
R;Nystedt, S.; Larsson, A.K.; Aberg, H.; Sundelin, J.
J. Biol. Chem. 270, 5950-5955, 1995
A;Title: The mouse proteinase-activated receptor-2 cDNA and gene. Molecular cloning and
A;Reference number: I48705; MUID:95197620; PMID:7890726
A;Accession: I48705
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-399 <RES>
A;Cross-references: EMBL:Z48043; NID:g663020; PIDN:CAA88097.1; PID:g663021
C;Superfamily: ATP receptor P2u

Query Match 27.28; Score 581.5; DB 2; Length 399;
Best Local Similarity 37.6%; Pred. No. 1.3e-38;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPISNIVTLWKLRLT-KSISLVIEHTNLATA 153
DB 63 IDEFSASILTKLTVPVPIIIVFVIGLPSNGMALWIFLTKKGPVAVIYMANLALA 122
QY 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRTTIVVYGNMYCAIILITCMGINRYLATAHP 213
DB 123 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182
QY 214 FTQKLPKRSFSLMCGIVMVFYMLPFVILKQEVHLVHSEITTCDDVDVDAESPSSF 273
DB 183 MGHPR-KKANIAGVSAIWLIFLTPVLYVMKQTIYIPALNITTCDDVLP-EVLVGD 240
QY 274 RFYVFSLAFPGFLIPPIVIFCVTTLI-----HKLKSKDRIMLGIYKAVILLIV 323
DB 241 MFNYFLSLAGVFLPALLTASAYVLMKILRSAMDEHSEKKRQA----IRLIITVLA 296
QY 324 IFTICFAPTNIIIVHANYHNTDSLYPMYLIALCLGSLNSCLDPFLYFVMSK 378
DB 297 MYFICFAPSLLLVHVFYLIKTQKQSHVYALVALCLSLNSCIDPFVYFVSK 351

RESULT 3
S66518
proteinase-activated receptor 2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Nov-2000
C;Accession: S66518; S64709; G02131
R;Nystedt, S.; Emlison, K.; Larsson, A.K.; Stroembeck, B.; Sundelin, J.
Eur. J. Biochem. 232, 84-89, 1995
A;Title: Molecular cloning and functional expression of the gene encoding the human pro-
A;Reference number: S66518; MUID:96048032; PMID:7556175
A;Accession: S66518
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <NYS>
A;Cross-references: EMBL:Z49993; NID:g1008084; PIDN:CAA90290.1; PID:g1008085
R;Boehm, S.K.; Kong, W.; Broemme, D.; Smeekens, S.P.; Anderson, D.C.; Connolly, A.; Kahn
Biochem. J. 314, 1009-1016, 1996
A;Title: Molecular cloning, expression and potential functions of the human proteinase-a
A;Reference number: S64709; MUID:96177879; PMID:8615752
A;Accession: S64709
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-137, 'A', 139-397 <BOE>
A;Cross-references: EMBL:U34038; NID:g1041728; PIDN:AB47871.1; PID:g1041729
A;Note: the authors translated the codon GTC for residue 68 as Ile and AAC for residue 2
R;Kahn, M.L.; Coughlin, S.R.
submitted to the EMBL Data Library, September 1995
A;Reference number: H00822
A;Accession: G02131
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 29-397 <KAH>
A;Cross-references: EMBL:U36753; NID:g1208539; PIDN:AAA90957.1; PID:g1208540
C;Genetics:

A;Map position: 5q13
A;Introns: 28/1
C;Superfamily: ATP receptor P2u
F;1-36/Domain: activation peptide #status predicted <APT>
F;1-25/Domain: signal sequence #status predicted <SIG>
F;37-397/Product: proteinase-activated receptor 2 #status predicted <MAT>

Query Match 26.6%; Score 569; DB 2; Length 397;
Best Local Similarity 32.7%; Pred. No. 1.3e-37;
Matches 120; Conservative 73; Mismatches 136; Indels 38; Gaps 8;

QY 22 LVAAGLLFLPVVTCQSGINVSNDNAKPTLTITIKSFNGQPONTPEEPFLSDIEGWTGATTTI 81
DB 9 LIGAALLAASLSCSGTIOGTRSSKGRSLICKVDG-----TSHV---TCKGVTV 55
QY 82 KAECPEDSISTUHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPISNIVTLWKLRLT-KS 140
DB 56 E-----TVPSVDFSAVLTGKLTITVFLPIVITVVFVGLPSNGMALWVFLFRTKKK 107
QY 141 ISLVIFHTNLATADLLFCVTLPPKIAHYHLNGNNWVGEVNCRTTIVVYGNMYCAIILIT 200
DB 108 HPVAYMANLADULLSVIWFPLKIAHYHLHNGNNWVGEALCNVLIGFFYGNMYCSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFSLMCGIVMVFYMLPFVILKQEVHLVHSEITTC 260
DB 168 CLSVQRYWVIVNPMGHSR-KKANIAIGISLAWLLILLVITPLVYVMKQTIYIPALNITTC 226
QY 261 HDVVDACESPSSF-----RFYFVSIAFFGFLIPFVILIFCYTTLTHKLS-----KDR 310
DB 227 HDVL-----PEQLLVGMDFNYFLSLAIGVFLPFAFLTASAYVLMIRMLRSSAMDENSEKK 281
QY 311 WLGYIKAVILLIVITFCFAPTNIIIVHANYHNTDSLYPMYLIALCLGSLNSCLDP 370
DB 282 RRKRAIKLVITVLAALICETPSNLLLVHVFYLIKTQKQSHVYALVALCLSLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 4
S17148
alpha-thrombin receptor - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C;Accession: S17148
R;Rasmussen, U.B.; Vouret-Craviari, V.; Jallat, S.; Schlesinger, Y.; Pages, G.; Paviran
FEBS Lett. 288, 123-128, 1991
A;Title: cDNA cloning and expression of a hamster alpha-thrombin receptor coupled to Ca
A;Reference number: S17148; MUID:91348247; PMID:1652467
A;Accession: S17148
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <RAS>
A;Cross-references: EMBL:X61958; NID:g940495; PIDN:CAA43957.1; PID:g49538
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 24.4%; Score 521; DB 2; Length 427;
Best Local Similarity 32.4%; Pred. No. 8.7e-34;
Matches 129; Conservative 78; Mismatches 125; Indels 66; Gaps 16;

QY 20 LILVAAG-----LLFLPVVTCQSGINVSNDNAKPTLTITIKSF---NGQPONTPEEPFLSD 70
DB 6 LLLVAAAGLSLCPGLLSRRVVPVQPESEMDATVNP-----RSFRLNPG-ENTFELIPLGD 60
QY 71 IEGWTGATTIKAECPEDSISTLH-----VNNATIGYLRSSLSQVIPAIIYLL 119
DB 61 EEEKNEST-----LPEGRAIYLNKSHSPALPFISSDASGVLTSPWLRLFIPSVVTVF 114
QY 120 FVVGVPISNIVTLWKLRLT-KSISLVIEHTNLATADLLFCVTLPPKIAHYHLNGNNWVFG 177
DB 115 FVVSPLNLITAIYAVFLVKMKVKK-PAVVYMLHLMADVLVSVLPKISYFSGSDMQFG 173

A:Gene: p2Y5
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein
F:15-40/Domain: transmembrane #status predicted <TM1>
F:51-74/Domain: transmembrane #status predicted <TM2>
F:89-109/Domain: transmembrane #status predicted <TM3>
F:133-153/Domain: transmembrane #status predicted <TM4>
F:177-201/Domain: transmembrane #status predicted <TM5>
F:227-248/Domain: transmembrane #status predicted <TM6>
F:269-292/Domain: transmembrane #status predicted <TM7>

Query Match 17.9%; Score 383; DB 2; Length 308;
Best Local Similarity 30.3%; Pred. No. 5.2e-23;
Matches 90; Conservative 59; Mismatches 104; Indels 44; Gaps 9;

QY 115 IYVLLFVVGPSNIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLFPKAYHLN 170
DB 20 VFSMVFLGLIANCAVAIYFTTLKVRNETTYML---NLASDLLFVTLFRI-YYPV 75

QY 171 GNNVFGVCMRITTVVFGNMYCAIILTCMGINRYLATAHPFTYQKLPKRSFSLMCG 230
DB 76 VRNWPFGDVLCKISVTLFTYNNYGSILFTLCISVDRFLAIVHPPRSKTLRTKRNARIVCV 135

QY 231 IYVWVFLYMLFPVILKQYHLVHSEITTCHEVDVADCESSPFRFYFVSLAFF----CF 286
DB 136 AVMITVLAGSTPASFFQSTNRQNNTEQRTCFE-----NFPSESTWTKYLSRIVIFTEIVGF 190

QY 287 LIPFVILIIICYTTLIHKLK-----SKRIWLGVIKAVLILVIFTICFAPTNIL 336
DB 191 FPIPLINVTCTWLTNLKPLTLNRKLSKKV-----LKMFLVHLVIFCFEVPNITL 246

QY 337 VIHANYYYHNTD-----SLYPMYLIALCLGSLNSCLDPFLFYFVMSKVVDQLN 384
DB 247 IL----YSLMRIQTWNCVWTAVRTWYPTVLCIAVSNCCFDPVIVYFTSDTNSELD 299

RESULT 11
ATP receptor P2u - mouse
C:Species: Mus musculus (house mouse)
C:Dates: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: A47556
R:Jusztig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.
Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993
A:Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.
A:Reference number: A47556; MUID:93281707; PMID:7685114
A:Accession: A47556
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <LUS>
A:Cross-references: GB:I14751; NID:g309457; PIDN:AAA39871.1; PID:g309458
C:Superfamily: ATP receptor P2u
C:Keywords: transmembrane protein

Query Match 17.7%; Score 378.5; DB 2; Length 373;
Best Local Similarity 26.9%; Pred. No. 1.4e-22;
Matches 100; Conservative 67; Mismatches 144; Indels 61; Gaps 13;

QY 69 SDIEGHTGATTIKACRPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYLLFVVGPSNI 128
DB 3 ADLEPW---NSTINGTWEGDELGYKCRFNEDFKYV-----LLPVSYGWCVLGLCLNV 52

QY 129 VTMLKLSLRTK---SISLVIFHTNLAIADLLFCVTLFPKAYHLNNGNMFGEVMCRIT 185
DB 53 VALYIFLCRLKTNASTTYWFH---LAVSDSLYASLPLIVYYVARGDHPFFSTVLCKLYR 110

QY 186 VFVFGNMYCAIILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIYVWVFLYMLP--- 242
DB 111 FLPTYNLCISILFTLCISVHRCGLVRLPLSLRWGRVARRVAVWVVLVLAQAPVLY 170

QY 243 FVILKQYHLVHSEITTCHEVDVADCESSPFRFYFVSLAFFGLFIPFVILIIICYTTLIH 302
DB 171 FVTTS-----VRGTRITCHD-TSARELFSHFVAYSSVMLGLL-FAVPFPSVILVYCYVLMAR 223

QY 303 KL-----KSKRIWLGVIKAVLILVIFTICFAPTNILVILHHA-----NYY 344
DB 224 RLLKPAYGTGGLPRAKRKS-----VRTIALVAVFALCFPLFPFHVTRTLVYSPRSLDLS 277

QY 345 YHNTDSLFPYMLIALCLGSLNSCLDPFLFYFVMSKV-----DQLNPKXSNMARP----- 392
DB 278 CHTLNAINMAYKITRPLASANSCLDPVLYFLAGQLRVFARDAKPPTEPTSPQARRKLG 337
QY 393 LKRPDRDIWEDI 404
DB 338 LHRPNRTVRKDL 349

RESULT 12
JC4737
G protein-coupled receptor P2Y1 - human
N:Alternate names: P2Y1 purinergic receptor; P2Y1 purinoceptor
C:Species: Homo sapiens (man)
C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: JC4737; JC4615; S54253
R:Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 221, 588-593, 1996
A:Title: Cloning and tissue distribution of the human P2Y1 receptor.
A:Reference number: JC4737; MUID:96205320; PMID:8630005
A:Accession: JC4737
A:Molecule type: DNA
A:Residues: 1-373 <GAN>
A:Cross-references: GB:S81950; NID:gl839438; PIDN:AA847091.1; PID:gl839439
R:Ayyanathan, K.; Webbs, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.P.
Biochem. Biophys. Res. Commun. 218, 783-788, 1996
A:Title: Cloning and chromosomal localization of the human P2Y1 purinoceptor.
A:Reference number: JC4615; MUID:96158962; PMID:8579591
A:Accession: JC4615
A:Molecule type: mRNA
A:Residues: 1-373 <AY>
A:Cross-references: GB:U42029; NID:gl147730; PIDN:AAA97872.1; PID:gl147731
A:Experimental source: erythro leukemia cells
R:Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.
submitted to the EMBL Data Library, May 1995
A:Description: Cloning of a human putative P2Y receptor.
A:Reference number: S54253
A:Accession: S54253
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-137,139-373 <LEO>
A:Cross-references: EMBL:249205; NID:g798835; PIDN:CAA89066.1; PID:g798836
C:Comment: This receptor belongs to a family of G protein-coupled receptors. It responds:
C:Genetics:
A:Gene: p2Y1; GDB:P2RY1
A:Cross-references: GDB:677125; OMIM:601167
A:Map position: 3pter-3qter
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane pro.
F:52-77/Domain: transmembrane #status predicted <TM1>
F:88-111/Domain: transmembrane #status predicted <TM2>
F:124-152/Domain: transmembrane #status predicted <TM3>
F:171-191/Domain: transmembrane #status predicted <TM4>
F:214-237/Domain: transmembrane #status predicted <TM5>
F:261-282/Domain: transmembrane #status predicted <TM6>
F:305-328/Domain: transmembrane #status predicted <TM7>
F:11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predic
F:330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predic
F:343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depe

Query Match 17.6%; Score 376.5; DB 2; Length 373;
Best Local Similarity 27.0%; Pred. No. 2.1e-22;
Matches 95; Conservative 72; Mismatches 130; Indels 55; Gaps 12;

QY 46 AKPTLTIKSFGNGPONTPEEPFLSDIEGTGATTIKACRPEDSISTLHVNNNA-TIGYLR 104
DB 8 AVPNGTDAFLAGPGSS-----WGNSTVA-----STAASVSSFKALTK 46

Query Match 16.7%; Score 357; DB 2; Length 380;
Best Local Similarity 27.9%; Pred. No. 7.4e-21;
Matches 100; Conservative 66; Mismatches 124; Indels 68; Gaps 12;

QY 52 IKSFNGGPQNT-----FEERPLSDIEGWTGATTITKAECPEDSISTLHVNN 97
DB 5 IQIFRGDPGPTCPSACLLPNSSWFPNWAESDSNGSVG-----EDQLESASHP 56

QY 98 ATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLKLSLRK---SISLVIFHTNLAIAD 154
DB 57 AI-----PVIITAVSVWFVGLVGNLSLWVFVIIRYTKMKTATNIYIF--NLALAD 105

QY 155 LLFCVTLPPKIAVHLNGNNWVGEVCMRTTIVVYGNMYCAIILTCMGINRYLATAPHP 214
DB 106 ALVTTMTMPQSAVYLM--NSWPFGLCKIVISIDYNNMFTSIPTLTMMSDRYIAVCHPV 164

QY 215 TYQ--KLKRSFSLMCGVVMVFLYMLPFVILKQEHVHLVHSEITTHCHDVVDACESPS 272
DB 165 KALDFRTPLKAKIINIC--IWLASSVGISAIVLGG-----TKVREDVDVIECLQ 213

QY 273 F-----RFYFVSLAFFGLIPFVILFICVTTLIHLKLS-----KDRIMWGYI 315
DB 214 FPDDEYSWMDLFWKICVVFAPVILVILVVCYTMILRLKSVRLSSGSRKDRNLRRIT 273

QY 316 KAVLLILVITFCFAPTNILVIHANYHNTDSLYFMYLIALCGLSGLNSCLDPLFY 373
DB 274 KLVVVVAVFIICTWPIHIFILVEALGSTSHSTAAL--SSYFYCIAGLYTNSSINPVLY 330

RESULT 16
A48227
kappa opioid receptor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jan-2000
C:Accession: A48227; J04138
R:Yasuda, K.; Raynor, K.; Kong, H.; Breder, C.D.; Takeda, J.; Reisine, T.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 90, 6736-6740, 1993
A:Title: Cloning and functional comparison of kappa and delta opioid receptors from mouse
A:Reference number: A48227; MUID:93342064; PMID:8393575
A:Accession: A48227
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <YAS>
A:Cross-references: GB:111065; NID:G348248; PIDN:AAA39363.1; PID:G348249
R:Liu, H.C.; Lu, S.; Augustin, L.B.; Felsheim, R.F.; Chen, H.C.; Loh, H.H.; Wei, L.N.
Biochem. Biophys. Res. Commun. 209, 639-647, 1995
A:Title: Cloning and promoter mapping of mouse kappa opioid receptor gene.
A:Reference number: J04138; MUID:95251663; PMID:7733933
A:Accession: J04138
A:Molecule type: mRNA
A:Residues: 1-380 <LIU>
A:Note: The authors translated the codon CAG for residue 365 as Glu
C:Comment: This receptor exists in different areas of the central and peripheral nervous
C:Genetics:
A:Gene: kor
C:Superfamily: vertebrate rhodopsin
C:Keywords: brain; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane

Query Match 16.7%; Score 357; DB 2; Length 380;
Best Local Similarity 27.9%; Pred. No. 7.4e-21;
Matches 100; Conservative 66; Mismatches 124; Indels 68; Gaps 12;

QY 52 IKSFNGGPQNT-----FEERPLSDIEGWTGATTITKAECPEDSISTLHVNN 97
DB 5 IQIFRGDPGPTCPSACLLPNSSWFPNWAESDSNGSVG-----EDQLESASHP 56

QY 98 ATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLKLSLRK---SISLVIFHTNLAIAD 154
DB 57 AI-----PVIITAVSVWFVGLVGNLSLWVFVIIRYTKMKTATNIYIF--NLALAD 105

QY 155 LLFCVTLPPKIAVHLNGNNWVGEVCMRTTIVVYGNMYCAIILTCMGINRYLATAPHP 214

DB 106 ALVTTMTMPQSAVYLM--NSWPFGLCKIVISIDYNNMFTSIPTLTMMSDRYIAVCHPV 164

QY 215 TYQ--KLKRSFSLMCGVVMVFLYMLPFVILKQEHVHLVHSEITTHCHDVVDACESPS 272

DB 165 KALDFRTPLKAKIINIC--IWLASSVGISAIVLGG-----TKVREDVDVIECLQ 213

QY 273 F-----RFYFVSLAFFGLIPFVILFICVTTLIHLKLS-----KDRIMWGYI 315

DB 214 FPDDEYSWMDLFWKICVVFAPVILVILVVCYTMILRLKSVRLSSGSRKDRNLRRIT 273

QY 316 KAVLLILVITFCFAPTNILVIHANYHNTDSLYFMYLIALCGLSGLNSCLDPLFY 373

DB 274 KLVVVVAVFIICTWPIHIFILVEALGSTSHSTAAL--SSYFYCIAGLYTNSSINPVLY 330

RESULT 17
T09508
intron 17 purinergic receptor P2Y5 - human
N:Alternate names: G-protein coupled receptor
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T09508
R:Bohm, S.K.; Trumpp, A.; Khitin, L.M.; Kong, W.; Pavan, D.G.; Bunnett, N.W.
submitted to the EMBL Data Library, April 1997
A:Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retinol
A:Reference number: Z16705
A:Accession: T09508
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <BOH>
A:Cross-references: EMBL:AF000546; NID:G22322068; PID:G22322069
C:Genetics:
A:Map position: 13
C:Superfamily: ATP receptor p2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.7%; Score 356; DB 2; Length 344;
Best Local Similarity 28.2%; Pred. No. 8e-21;
Matches 87; Conservative 72; Mismatches 110; Indels 40; Gaps 11;

QY 97 NATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAI 152

DB 5 NSSHCIFYNDSFKYLYGCMFSWVFLGLVSNCAIIFIICVLKVRNETTYMI---NLAM 61

QY 153 ADLLFCVTLPPKIAIYHLNGNNWVGEVCMRTTIVVYGNMYCAIILTCMGINRYLATAP 212

DB 62 SDLLFVTLPPRI--FYFTRNWPFGDLLCKISVLMFYTNMYSILFLTCISVDRFLAIVY 120

QY 213 PTYQKLPKRSFSLMCGVVMVFLYMLPFVILKQEHVHLVHSEITTHCHDVVDAC--ESP 270

DB 121 PFKSKTLRTKNAKIVCTGVMLTVIGGAPAVFVQS---THSQ---GNNASEACFENFP 173

QY 271 SFPRFYFVSLAFF---GFLIPFVILFICVTTLIHLK-----SKDRIMWGYIK 316

DB 174 EATWTKYLSRIVFIEIVGFFIPLINVTCSMVLKTLTKPVTLRSRKINKTKV----LK 229

QY 317 AVLLILVITFCFAPTNILVIH--ANYVYHNTD---SLYFMYLIALCGLSGLNSCLDPP 371

DB 230 MIFVHLIIFCFCEVFPYNNILYLSLVRTQTFVNCVVAARVMTYPTILCIAVSNCCFDP 289

QY 372 LYFVMSKV 380

DB 290 VYVFTSDTI 298

RESULT 18
JN0694
angiotensin II receptor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: JN0694; S47627
R:Ji, H.; Sandberg, K.; Zhang, Y.; Catt, K.J.
Biochem. Biophys. Res. Commun. 194, 756-762, 1993

A:Title: Molecular cloning, sequencing and functional expression of an amphibian angiotensin receptor
A:Reference number: JN0694; MUID:93343933; PMID:7688227
A:Accession: JN0694
A:Molecule type: mRNA
A:Residues: 1-362 <JH>
A:Cross-references: GB:116463; NID:9387890; PIDN:AAA9647.1; PID:9387891
R:Nishimatsu, S.; Koyasu, N.; Sugaya, T.; Ohnishi, J.; Yamagishi, K.; Miya
Biochim. Biophys. Acta 1218, 401-407, 1994
A:Title: Isolation and characterization of two alternatively spliced complementary DNAs
A:Reference number: S47627; MUID:94325348; PMID:7519446
A:Accession: S47627
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-362 <NIS>
A:Cross-references: GB:S73274; NID:G625009; PIDN:AAC60749.1; PID:G625010
C:Superfamily: vertebrate rhodopsin
C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F:34-54/Domain: transmembrane #status predicted <TM1>
F:66-86/Domain: transmembrane #status predicted <TM2>
F:105-125/Domain: transmembrane #status predicted <TM3>
F:146-166/Domain: transmembrane #status predicted <TM4>
F:192-212/Domain: transmembrane #status predicted <TM5>
F:241-260/Domain: transmembrane #status predicted <TM6>
F:285-305/Domain: transmembrane #status predicted <TM7>
F:318-343/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:347/Binding site: phosphate (Thr) (covalent) #status predicted
Query Match 16.6%; Score 355.5; DB 2; Length 362;
Best Local Similarity 29.1%; Pred. No. 9.2e-21;
Matches 92; Conservative 56; Mismatches 127; Indels 41; Gaps 9;
QY 89 SISTLVHNNATIGYLRSSLSSTQV---IPAIYLLFVVGVPNS--IVTLWKLISRTKLSISL 143
DB 6 TVETSDVERIANCVNSGSHNYIFIAPIIYSTIFVVGFGNSMVVIYISYNNKMKTV- 64
QY 144 VIFHTNLAIADLLFCVLPFKIAYHLNGNNWFEVGMCRITTVFYGNNMYCAILILTCMG 203
DB 65 SIFLMNLALSDLCFVITPLWAAATYAMHYHWPFGNGLCKVASTAITNLNLTIVTLTCLLS 124
QY 204 INRYLATAPFTYQKLPKRSFSLMCGIYVWVFLYMLPFVILKQYHLVHSEITTCDDV 263
DB 125 IDRYSAIVHPMKSRIRWRTAMVARLTCVGIWLAFLASMPISIIYRQIYLFHDTNQTVCAIV 184
QY 264 VDACESPSRFRFYFVSLA---FFGFLIPFVILIFCYTTLIHKLKAS-----KDRW 311
DB 185 YD-----SGHIYFMVGMSLAKNIVGFLIPFLILITSLTGKTLKEVYRAQRNDDIF 238
QY 312 LGYIKAVLLILVITTCFAPTNIIL-----VIHANYYYHNTDSLYFMYLIALCLG 362
DB 239 ----KMIVAVLLFFFCWIPYQVFTFLDLVLIQMDVIONCKWY----DIVDTGMPITICIA 290
QY 363 SLNSCLDPLFYFVMSK 378
DB 291 YFNSCLNPLFYGFPGK 306
RESULT 19
I38435
angiotensin receptor homolog APJ - human
C:Species: Homo sapiens (man)
C:Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 21-Jul-2000
C:Accession: I38435
R:O'Dowd, B.F.; Heiber, M.; Chan, A.; Heng, H.H.; Tsui, L.; Kennedy, J.L.; Shi, X.; Petr
Gene 136, 355-360, 1993
A:Title: A human gene that shows identity with the gene encoding the angiotensin receptor
A:Reference number: I38435; MUID:94124031; PMID:8294032
A:Accession: I38435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-380 <RES>
A:Cross-references: EMBL:U03642; NID:9425351; PIDN:AAA18954.1; PID:9425352
C:Genetics:

A:Gene: APJ
A:Map position: 11q12
A:Introns: #status absent
C:Superfamily: vertebrate rhodopsin
Query Match 16.6%; Score 355; DB 2; Length 380;
Best Local Similarity 27.3%; Pred. No. 1.1e-20;
Matches 89; Conservative 66; Mismatches 127; Indels 44; Gaps 9;
QY 70 DIEGTGATTIKACPCPDSISTLVHNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIV 129
DB 6 DFDNYTGADN--QSECE-----YTDKSSGALIPAIYMLVFLTGTTGNGL 48
QY 130 TLWKL--SLRTKISLVIFHTNLAIADLLFCVLPFKIAYHLNGNNWFEVGMCRITTV 187
DB 49 VLTWTFRSRKRKRADIFIASLAVADLTFTVTLPLWATYTYRDYDNWFGTFCKLSVL 108
QY 188 FYGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIYVWVFLYMLPFVILK 247
DB 109 IFVNNYASVFCITGLSDFRYLAIVRPVANARLRLRVSGAVATAVLWLAALLAMPVWL 168
QY 248 QEYHLVHSEITTC---HDVDACESPSRFRFYFVSLAFFGFLIPFVILIFCYTTL---- 300
DB 169 TTGDLNTTKVQCYNDYSMTATVSEWAVEVGLGVSTTVGVFPFTIMLTCTCYFFIAQT 228
QY 301 -----IHKLKSODRIWLGVIKAVLLILVITTCFAPTNIILVIHANYYYH---NTD 349
DB 229 AGHPRKERIEGLRKRRRL-----LSIIVLVITFALCWMPYHLVLYMLGSLHWPCCDF 284
QY 350 SLYFMYLII--ALCLGSLNSCLDPLFY 373
DB 285 -LFLMNIFFPYCTCISYVNSCLNPLFY 309
RESULT 20
JC2338
kappa opioid receptor 1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 19-May-2000
C:Accession: JC2338; A55354; I57005; G01546
R:Mansson, E.; Bare, L.; Yang, D.
Biochem. Biophys. Res. Commun. 202, 1431-1437, 1994
A:Title: Isolation of a human kappa opioid receptor cDNA from placenta.
A:Reference number: JC2338; MUID:94338360; PMID:8060324
A:Accession: JC2338
A:Molecule type: mRNA
A:Residues: 1-380 <MAN>
A:Cross-references: GB:U11053; NID:9532059; PIDN:AAA20985.1; PID:9532060
A:Experimental source: placenta
R:Wang, J.B.; Johnson, P.S.; Wu, J.M.; Wang, W.F.; Uhl, G.R.
J. Biol. Chem. 269, 25966-25969, 1994
A:Title: Human kappa opiate receptor second extracellular loop elevates dynorphin's affi
A:Reference number: A55354; MUID:95014415; PMID:7929306
A:Accession: A55354
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 136-279 <WAN>
A:Cross-references: GB:L36130; NID:9598184; PIDN:AAA63646.1; PID:9598185
R:Zhu, J.; Chen, C.; Xue, J.
Life Sci. 56, 201-207, 1995
A:Title: Cloning of a human kappa opioid receptor from the brain.
A:Reference number: I57005
A:Accession: I57005
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1,'D',3-380 <ZHU>
A:Cross-references: GB:L37362; NID:g722617; PIDN:AAA63906.1; PID:g722618
R:Grandy, D.K.
submitted to the EMBL Data Library, November 1994
A:Reference number: G07718
A:Accession: G01546
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 132-203 <GRA>
A:Cross-references: EMBL:U16860; NID:g595932; PIDN:AAA56758.1; PID:g595933
C:Comment: This receptor preferentially binds dynorphins.
C:Genetics:
A:Gene: GDB:OPR1; KOR
A:Cross-references: GDB:132651; OMIM:165196
A:Map position: 8q11.2-8q11.2
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; receptor; transmembrane protein
F:60-85/Domain: transmembrane #status predicted <TM1>
F:95-114/Domain: transmembrane #status predicted <TM2>
F:133-154/Domain: transmembrane #status predicted <TM3>
F:177-199/Domain: transmembrane #status predicted <TM4>
F:228-251/Domain: transmembrane #status predicted <TM5>
F:275-296/Domain: transmembrane #status predicted <TM6>
F:311-333/Domain: transmembrane #status predicted <TM7>

Query Match 16.5%; Score 353.5; DB 2; Length 380;
Best Local Similarity 27.9%; Pred. No. 1.4e-20;
Matches 94; Conservative 66; Mismatches 126; Indels 51; Gaps 11;

Qy 59 PONTPEEFPLSDIEGWTGATTTTIKAECPEDSISTLHVNNATIGYLRSSLSQTQVIPAIYIL 118
Db 23 PNSSAWFP-----GWAEPDSNGSAGEDAQLEPAHISPAI-----PVIITAVYSV 68

Qy 119 LFVWGVPNSIVTLWLKLSLRTK---SISLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNNVV 175
Db 69 VFVWGLVGNLSVMFVIIRYTKMTATNIYIF--NLALADALVTTTTPQSTVYLM-NSWP 125

Qy 176 FGEVMCRITTVFVGYGMYCAILLTCGINRYLATAHPFTYQ--KLPKRSFSLMCGIIVW 233
Db 126 FGDVLCKIVISIDYNNFTSIFLTMTMSVDRIYAVCHPVKALDPTPLKAKIINIC--IW 183

Qy 234 VMVFLYMLFPVILKQEVHLVHSEITTCDDVDACESPSSF-----REFYFVSLAPFG 285
Db 184 LSSSVGISAIVLGG-----TKREDVDVIECSLQPPDDYSWDLFMKICVFIPA 234

Qy 286 FLIPFVIIIFCYTTLIHKLKS-----KDIRLWGLYIKAVLLILVIFTICFAPTNIIL 336
Db 235 FVIPLLIIIVCYTLMILRLKS VRLSSREKDRNLRIITRLVLVVAVVAVVVCWTPIHIFI 294

Qy 337 VIHHANYHYNTDSLFWYLIALLCLGSLNSCLDPFLY 373
Db 295 LVEALGSTSHSTAAL--SSYFICIALGYTNSLSNPILY 330

RESULT 21

kappa opioid receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Dec-1993 #sequence revision 19-Oct-1995 #text_change 20-Jun-2000
C:Accession: S36143; S38825; S36102; S39015; A48789
R:Nishi, M.; Takeshima, H.; Fukuda, K.; Kato, S.; Mori, K.
FEBS Lett. 330, 77-80, 1993

A:Title: cDNA cloning and pharmacological characterization of an opioid receptor with hi
A:Reference number: S36143; MUID:93380575; PMID:8396539
A:Accession: S36143
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <NIS>
A:Cross-references: GB:D16534; NID:g409390; PIDN:BAA03971.1; PID:g415310
R:Chen, Y.; Mestek, A.; Liu, J.; Yu, L.
Biochem. J. 295, 625-628, 1993

A:Title: Molecular cloning of a rat kappa opioid receptor reveals sequence similarities
A:Reference number: S38825; MUID:94059008; PMID:8240267
A:Accession: S38825
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <CHE>
A:Cross-references: GB:I22001; NID:g409236; PIDN:AAA41495.1; PID:g409237
R:Minami, M.; Toyota, T.; Kacao, Y.; Maekawa, K.; Nakamura, S.; Onogi, T.; Kaneko, S.; Sat
FEBS Lett. 329, 291-295, 1993

A:Title: Cloning and expression of a cDNA for the rat kappa-opioid receptor.

Query Match 16.4%; Score 350; DB 2; Length 380;
Best Local Similarity 27.6%; Pred. No. 2.6e-20;
Matches 96; Conservative 65; Mismatches 131; Indels 56; Gaps 11;

QY 58 GQNTFEERPLSD-----IGWTGATTIKACCPEDSISILHVNNAIGYLRSL 107
DB 7 GPAQASELPARNACLLPNAGSAPGQDEQLEPAHISPAI-----58

QY 108 STQVPAIYLLFVVGVPNSNIVTLWKLRLTK--SISLVIFHTNLAIAADLLFCVTLPPK 164
DB 59 -PVIITAVSVFVVGVLGNSLVMFVIIRYTKMTATNIYIF--NLAJADALVTTMTPEQ 115

QY 165 IAYHLNGNNWFGVEMCRITTVVFGNMYCAILILTCMGINRYLATAPHTTYQ--KLPRK 222
DB 116 STVYLM-NSWPGDVLCKIVISIDYNNMFTSIFLTMTMSVDRIYVACHPVKALDFRTPLK 174

QY 223 SFSLMCGIWMVFLYMLPFVILKQEHVHSEITTCDDVDVACESPSSP-----R 274
DB 175 AKIINIC--IWLSSSVGISAILGG-----TKREDVDIIECSLQFPDDDDYSWMD 223

QY 275 FYFVSLAFEGFLPVIIFVCTTLIHLKLS-----KDRWLGVYIKAVLLILVIF 325
DB 224 LFMKICVFVFAFVPIVLIIVCYTLMLRLKSVRLSSGREGKDRNLRRITRLVLVVAVF 283

QY 326 TICFAPTNILVIHANYYYHNTDSLYFMYLIACLSLNSCLDPPLY 373
DB 284 IICWTPIHIFILVEALGSGSHSTAAL--SSYFICIALGYTNSSINPLY 330

RESULT 23
A54946
P-2U nucleotide receptor - human
C:Species: Homo sapiens (man)
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 17-Mar-1999
C:Accession: A54946
R:Parf, C.E.; Sullivan, D.M.; Paradiso, A.M.; Lazarowski, E.R.; Burch, L.H.; Olsen, J.C.
Proc. Natl. Acad. Sci. U.S.A. 91, 3275-3279, 1994
A:Title: Cloning and expression of a human P-2U nucleotide receptor, a target for cystic
A:Reference number: A54946; MUID:94211846; PMID:8159738
A:Accession: A54946
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-375 <PAR>
A:Cross-references: GB:U07225
A>Note: parts of this sequence were confirmed by protein sequencing
C:Genetics:
A:Gene: GDB:P2R2; HP2U; P2U
A:Cross-references: GDB:362713; OMIM:600041
A:Map position: 11q13.5-11q14.1
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.3%; Score 347.5; DB 2; Length 375;
Best Local Similarity 28.7%; Pred. No. 4.1e-20;
Matches 83; Conservative 57; Mismatches 108; Indels 41; Gaps 10;

QY 111 VIPAIYLLFVVGVPNSNIVTLWKLRLTK--SISLVIFHTNLAIAADLLFCVTLPEKIA 167
DB 35 LLPVSVGVVGVGLCLNLAVGLYIFLCRLKTNASTYMFH--LAVSDALYAASLPVLVY 92

QY 168 HUNGNNWFGVEMCRITTVVFGNMYCAILILTCMGINRYLATAPHTTYQKLPKRSFSL 227
DB 93 YARGDHWPFSTVLCKLVRLFYTNLYCSILFLTCISVHRLGLVRLSLRWGRARYARR 152

QY 228 MCGIWMVFLYMLP---FVILKQEHVHSEITTCDDVDVACESPSSPFRFYFVSLAPF 284
DB 153 VAGAVWVLVLAQAPVLYFVTTTSARGPL-----TCHD-TSAPELFSRFVAYSSVMLGL 205

QY 285 GELIPVLIIFCVCTTLIHLK-----KSKDRWLGVYIKAVLLILVIFITCF 330
DB 206 -FAVPFAVLVLCVWARRLLKPAYTSGGLPRAKRS-----VRTIAVLAVFALCEL 258

QY 331 PTNIIILVIHHA-----NYYHNTDSLYFMYLIACLSLNSCLDPPLYFV 375

DB 259 PFHVTRTLYYFRSLDLSCHTLNAINMAYKVR-LASANSCLDPVLYEL 306
RESULT 24
B57641
G protein-coupled receptor 4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 29-Sep-1999
C:Accession: B57641
R:Mahadevan, M.S.; Baird, S.; Bailly, J.E.; Shutler, G.G.; Sabourin, L.A.; Tsilfidis, C.
Genomics 30, 84-88, 1995
A:Title: Isolation of a novel G protein-coupled receptor (GPR4) localized to chromosome
A:Reference number: A57641; MUID:96129306; PMID:8595909
A:Accession: B57641
A>Status: preliminary; nucleic acid sequence not shown; translation not shown; significance
A:Molecule type: DNA
A:Residues: 1-362 <MAH>
A:Cross-references: GB:U22108; NID:G722282; PIDN:AAA98458.1; PID:G722283
C:Genetics:
A:Introns: #status absent
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

Query Match 15.6%; Score 333.5; DB 2; Length 362;
Best Local Similarity 28.2%; Pred. No. 5.1e-19;
Matches 81; Conservative 60; Mismatches 109; Indels 37; Gaps 7;

QY 113 PAIYILLFVVGVPNSNIVTLWKL--SLRTKSTSLVIFHTNLAIAADLLFCVTLPEKIA 170
DB 21 PSLYIFVIGVGLPTNCLRLMAAYRQVRQNR-ELGVYLMNLSIADLLYICTLPLWVDYFLH 79

QY 171 GNNVFGVEMCRITTVVFGNMYCAILILTCMGINRYLATAPHTTYQKLPKRSFSLMCG 230
DB 80 HDNTHGPGSKCLFGFIFTYTHIYISIAFLCCISVDRIYLAHPLRFALRRVKTAVAVSS 139

QY 231 IWMVFLYMLPFFVILKQEHVHSEITTCDDVDVACESPSSPFRFYF-----VS 280
DB 140 VVWAT-----ELGANSVPLFHDDELFRDYNHTFCFEKFPMEGWVAMNLY 184

QY 281 LAFFGFLIPFVLIIFCVCTTLIHLK---SKDRWLGVYIKAVLLILV-IFTICFAPTNIL 336
DB 185 RVFVGFLFPWMLLSYRGILRAVGSVSTERQEKAKIKRLALSIAIVLVCFAPYHLL 244

QY 337 VIHANYYYHNTD-----SLVFMYLIALCLGSLNSCLDPPLYFVMSK 378
DB 245 LRSRAVYLGHWPDCGFEEVFSVAVHSSLAFTSLNLCVADPILYCIUNE 291

RESULT 25
I55450
G protein-coupled P2 receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C:Accession: I55450
R:Chang, K.; Hanaoka, K.; Kumada, M.; Takuwa, Y.
J. Biol. Chem. 270, 26152-26158, 1995
A:Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.
A:Reference number: I55450; MUID:96064682; PMID:7592819
A:Accession: I55450
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <RES>
A:Cross-references: GB:D63665; NID:G1066007; PIDN:BAA09816.1; PID:G1066008
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor

Query Match 15.6%; Score 333; DB 2; Length 328;
Best Local Similarity 27.9%; Pred. No. 5e-19;
Matches 87; Conservative 54; Mismatches 109; Indels 62; Gaps 12;

QY 104 RSSLSTQVPIPIYILLFVVGVPNSNIVTLWKL--SLRTKSTSLVIFHTNLAIAADLLFCVTL 161
DB 104 RSSLSTQVPIPIYILLFVVGVPNSNIVTLWKL--SLRTKSTSLVIFHTNLAIAADLLFCVTL 161

Db 21 REDFKLLPPVYVVLVGLPLNVCVIAQICASRRLTRS-AVYTLNLALADLLVACSL 79
QY 162 PFKIAVHLGNWVFGVCMCRITTVVFGYGMVCAIILITCMGIRNRYLATAHPPT-YOKLP 220
Db 80 PLLIYNYARGDHPFGDLACRLVRFLFYANLHGSILFLICISFORVLGICHPLAPHHKRG 139
QY 221 KRSFSLMCGIYVMVFLYMLPVILKQEVHLVHSEITTCCHDVVDACESPS--SFRFY-Y 277
Db 140 GRBAWVVCVWVWVTAQCLPTAFAATG--IQNRNRYCYDL-----SPPILSTRYLPY 192
QY 278 FVSLAFPGELIPVILFCYTTLIHLKSKDRILWLG-----YIKAVLLILVIF 325
Db 193 GMALTIGFLPPTALLACVCRMARLRCQD-----GPAGVQARRSKAARWVVAAYF 248
QY 326 TICFAPNTIILVIHANYYYHNTDSLYF-----MYLIALCLGSLNSC 367
Db 249 VISFLP-----PHIKTAYLAVRSTPGVSCPVEITFAAAYKGTFRPFASANSV 295
QY 368 LDPFL-YFVNSK 378
Db 296 LDPILFYFTQOK 307

RESULT 26

IS1372
angiotensin II receptor - turkey
C:Species: Meleagris gallopavo (common turkey)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Nov-1999
C:Accession: IS1372; P00449
R:Murphy, T.J.; Nakamura, Y.; Takeuchi, K.; Alexander, R.W.
Mol. Pharmacol. 44, 1-7, 1993
A:Title: A cloned angiotensin receptor isoform from the turkey adrenal gland is pharmacologically identical to the human receptor
A:Reference number: IS1372; MUID:93341466; PMID:8341266
A:Accession: IS1372
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-359 <MUR>
A:Cross-references: GB:I23203; NID:G349735; PIDN:AAA03560.1; PID:G349736
R:Garsia, R.V.; McIlroy, P.J.; Kowaleki, K.I.; Tilly, J.L.
Biochem. Biophys. Res. Commun. 191, 1073-1080, 1993
A:Title: Isolation of turkey adrenocortical cell angiotensin II (AII) receptor partial cDNA
A:Reference number: P00449; MUID:93221469; PMID:7916599
A:Accession: P00449
A:Molecule type: mRNA
A:Residues: 53-91, 'C', 93-162, 'SSFTIVY', 171-291 <CAR>
A:Cross-references: GB:S58041; NID:G299373; PIDN:AAB26041.1; PID:G299374
A:Experimental source: liver
C:Superfamily: vertebrate rhodopsin
C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F:53-64/Domain: intracellular #status predicted <IN1>
F:53-64/Domain: transmembrane #status predicted <TM1>
F:137-102/Domain: extracellular #status predicted <EX1>
F:103-124/Domain: transmembrane #status predicted <TM2>
F:125-142/Domain: intracellular #status predicted <IN2>
F:143-162/Domain: transmembrane #status predicted <TM3>
F:163-192/Domain: extracellular #status predicted <EX2>
F:193-214/Domain: transmembrane #status predicted <TM4>
F:215-238/Domain: intracellular #status predicted <IN3>
F:239-262/Domain: transmembrane #status predicted <TM5>
F:263-274/Domain: extracellular #status predicted <EX3>
F:275-291/Domain: transmembrane #status predicted <TM6>
F:136/Binding site: phosphate (Ser) (covalent) #status predicted
F:141,233/Binding site: phosphate (Thr) (covalent) #status predicted
F:176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.5%; Score 332; DB 2; Length 359;
Best Local Similarity 26.9%; Pred. No. 6.6e-19;
Matches 82; Conservative 65; Mismatches 140; Indels 18; Gaps 6;

QY 87 EDSISTLHVNATIGYLRSLSTQVIPAIIYLLFVGVGPSN--IVTLWLKLSLTGKSLV 144
Db 8 EETVKRIHDPVSG--RHSYIYIMVPTVYSIIIFIIGNLSLWVIYCYMKLKTVA-S 64

QY 145 IFHTNLAIADLFCVTLFPFKIAYHLGNWVFGVCMCRITTVVFGYGMVCAIILITCMGI 204
Db 65 IFLLNLALADLFCVTLFPFKIAYHLGNWVFGVCMCRITTVVFGYGMVCAIILITCMGI 124
QY 205 NRYLATAHPPTYOKLPKRSFSLMCGIYVMVFLYMLPVILKQEVHLVHSEITTCCHDVV 264
Db 125 DRYLAIVHPVKSIRRTTMPVARVTCIVILWLAGVASLPVTHRNIFFAENLNMVTCGFRY 184
QY 265 DACESPSFRFYVYVSLAFFGLIPFVIFIIFCYTTLIHLKSKDRILWLG-----YIKAV 318
Db 185 D--NNNTLRVGLGSLKNLGLFPLIILSYTLNWKTKKAYQIQNRNTRDDIFKMI 242
QY 319 LLILVIFTCFAPTNI-----ILVTHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPPFLY 373
Db 243 VAIVFFFFSWIPHQVFTFLDLVLIQLHVITDCKITDIVDTAMPFTICIAVFNCLNPFFY 302
QY 374 FVNSK 378
Db 303 VFFGK 307

RESULT 27

JC2134
angiotensin II receptor type 1A - rat
N:Alternate names: AT1a receptor; AT1 receptor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 24-Nov-1999
C:Accession: JC2134; S15404; S20424; J01055
R:Conchon, S.; Monnot, C.; Sirieix, M.E.; Bihoreau, C.; Corvol, P.; Clauser, E.
Biochem. Biophys. Res. Commun. 199, 1347-1354, 1994
A:Title: Synthetic cDNA encoding the rat AT1a receptor: a useful tool for structure-function studies
A:Reference number: JC2134; MUID:94197726; PMID:8147879
A:Accession: JC2134
A:Molecule type: mRNA
A:Residues: 1-359 <CON>
A:Note: The amino acid sequence of this protein is not given
R:Murphy, T.J.; Alexander, R.W.; Griendling, K.K.; Runge, M.S.; Bernstein, K.E.
Nature 351, 233-236, 1991
A:Title: Isolation of a cDNA encoding the vascular type-1 angiotensin II receptor.
A:Reference number: S15404; MUID:91251901; PMID:2041570
A:Accession: S15404
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-359 <MUR>
A:Cross-references: GB:X62295; NID:G57773; PIDN:CAA44183.1; PID:G57774
R:Iwai, N.; Inagami, T.
FEBS Lett. 238, 257-260, 1992
A:Title: Identification of two subtypes in the rat type I angiotensin II receptor.
A:Reference number: S20423; MUID:92183879; PMID:1544458
A:Accession: S20424
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-359 <IWA>
R:Iwai, N.; Yamano, Y.; Chaki, S.; Konishi, F.; Bardhan, S.; Tibbette, C.; Sasaki, K.; I
Biochem. Biophys. Res. Commun. 177, 299-304, 1991
A:Title: Rat angiotensin II receptor: cDNA sequence and regulation of the gene expression
A:Reference number: JQ1055; MUID:91254291; PMID:2043116
A:Accession: JQ1055
A:Molecule type: mRNA
A:Residues: 1-80, 'C', 82-108, 'T', 110-359 <IW2>
A:Cross-references: GB:M74054; NID:G202918; PIDN:AAA40738.1; PID:G202919
A:Experimental source: kidney
C:Comment: AT1 receptor consists of two closely related AT1 isoforms of angiotensin II I
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; GTP binding; receptor; transmembr
F:27-86/Domain: transmembrane #status predicted <TM1>
F:103-123/Domain: transmembrane #status predicted <TM3>
F:147-167/Domain: transmembrane #status predicted <TM4>
F:195-218/Domain: transmembrane #status predicted <TM5>
F:240-263/Domain: transmembrane #status predicted <TM6>
F:278-299/Domain: transmembrane #status predicted <TM7>
F:4,176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.3%; Score 327; DB 2; Length 359;
Best Local Similarity 28.0%; Pred. No. 1.6e-18;
Matches 90; Conservative 51; Mismatches 129; Indels 52; Gaps 8;

QY 87 EDISITLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVSNN-----IVTLWKLRLRKSIS 142
DB 8 EDIGKRIQDDCPKAG--RHISYIFWMLPTLSIIFVVGIFGNSLVVIVFYMKLKTVA-- 63

QY 143 LVIFHTNLAIADLFCVTLTPFKIAYHLNGNNWVFGVMCRITTVVFGYNNYCAILLITCM 202
DB 64 -SVFLNLALADLCFLTLPLMAVYTAMEYRWPFGNHLCKIASASVFNLYASVFLTLCL 122

QY 203 GINRYLATAPFTYQKLPKRSFSLMLCGIWMVFLYMLPFVILKOEYHLVHSEITTCND 262
DB 123 SIDRYLAIVHPMKSLRLRRTMLVAKVTCIIILWLAGLASLPAVHRNVYFIENITIVC-- 180

QY 263 VVDACESPSSFRFY-----FVSLAPFGFLIPFVIIIFCYTTLIHLKSKDRIW 311
DB 181 -----AFHYESRNSLTPIGLTKNIGLGFLLFLLTSTYLLWKLKAYEIQ 229

QY 312 LG-----YIKAVLLILVITFCAPTNIIL-----VIHANYYYHNTDSLYFMYL 356
DB 230 KNKPRNDDIFRIIMAILVFFFSWVPHQIFTFDLVLIQLGVHDCIKI-----SDIVDTAMP 285

QY 357 IALCLGSLNSCLDPFLYFVMSK 378
DB 286 ITICIAVFNCLNPLFYGLGK 307

RESULT 28
JH0621
angiotensin II receptor 1A - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 24-Nov-1999
C:Accession: JH0621; JCI193
R:Sasamura, H.; Hein, L.; Krieger, J.E.; Pratt, R.E.; Kobilka, B.K.; Dzau, V.J.
Biochem. Biophys. Res. Commun. 185, 253-259, 1992
A:Title: Cloning, characterization, and expression of two angiotensin receptor (AT-1) is
A:Reference number: JH0621; MUID:92287102; PMID:1599461
A:Accession: JH0621
A:Molecule type: DNA
A:Residues: 1-359 <SAS>
A:Cross-references: GB:S37484; NID:G249945; PIDN:AAB22269.1; PID:G249946
A:Experimental source: strain Balb/c
R:Yoshida, H.; Kakuchi, J.; Guo, D.F.; Furuta, H.; Iwai, N.; van der Meer-de Jong, R.; I
Biochem. Biophys. Res. Commun. 186, 1042-1049, 1992
A:Title: Analysis of the evolution of angiotensin II type 1 receptor gene in mammals (m
A:Reference number: JCI193; MUID:92359981; PMID:1497638
A:Accession: JCI193
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-6,'I',8-19,'IS',22-37,'M',39-133,'K',135-359 <YOS>
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:37-57/Domain: transmembrane #status predicted <TM1>
F:65-85/Domain: transmembrane #status predicted <TM2>
F:103-123/Domain: transmembrane #status predicted <TM3>
F:145-165/Domain: transmembrane #status predicted <TM4>
F:201-220/Domain: transmembrane #status predicted <TM5>
F:241-261/Domain: transmembrane #status predicted <TM6>
F:286-306/Domain: transmembrane #status predicted <TM7>
F:4,176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:331,338/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 15.3%; Score 326; DB 2; Length 359;
Best Local Similarity 28.0%; Pred. No. 2e-18;
Matches 90; Conservative 50; Mismatches 130; Indels 52; Gaps 8;

QY 87 EDISITLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVSNN-----IVTLWKLRLRKSIS 142
DB 8 EDIGKRIQDDCPKAG--RHISYIFWMLPTLSIIFVVGIFGNSLVVIVFYMKLKTVA-- 63

QY 143 LVIFHTNLAIADLFCVTLTPFKIAYHLNGNNWVFGVMCRITTVVFGYNNYCAILLITCM 202

DB 64 -SVFLNLALADLCFLTLPLMAVYTAMEYRWPFGNHLCKIASASVFNLYASVFLTLCL 122

QY 203 GINRYLATAPFTYQKLPKRSFSLMLCGIWMVFLYMLPFVILKOEYHLVHSEITTCND 262

DB 123 SIDRYLAIVHPMKSLRLRRTMLVAKVTCIIILWLAGLASLPAVHRNVYFIENITIVC-- 180

QY 263 VVDACESPSSFRFY-----FVSLAPFGFLIPFVIIIFCYTTLIHLKSKDRIW 311

DB 181 -----AFHYESRNSLTPIGLTKNIGLGFLLFLLTSTYLLWKLKAYEIQ 229

QY 312 LG-----YIKAVLLILVITFCAPTNIIL-----VIHANYYYHNTDSLYFMYL 356

DB 230 KNKPRNDDIFRIIMAILVFFFSWVPHQIFTFDLVLIQLGVHDCIKI-----ADIVDTAMP 285

QY 357 IALCLGSLNSCLDPFLYFVMSK 378

DB 286 ITICIAVFNCLNPLFYGLGK 307

RESULT 29
JC2492
G protein-coupled receptor 1 - rat
N:Alternate names: GPR-1
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 08-Oct-1999
C:Accession: JC2492
R:Marchese, A.; Cheng, R.; Lee, M.C.; Porter, C.A.; Heiber, M.; Goodman, M.; George, S.I
Biochem. Biophys. Res. Commun. 205, 1952-1958, 1994
A:Title: Mapping studies of two G protein-coupled receptor genes: An amino acid differer
A:Reference number: JC2492; MUID:95110347; PMID:7811287
A:Accession: JC2492
A:Molecule type: mRNA
A:Residues: 1-353 <MAR>
A:Cross-references: GB:S74702; NID:G786483; PIDN:AAB322978.1; PID:G786484
C:Superfamily: vertebrate rhodopsin
C:Keywords: glycoprotein; lipoprotein; phosphoprotein; receptor; thiolester bond; trans
F:74-94/Domain: transmembrane #status predicted <TM2>
F:112-133/Domain: transmembrane #status predicted <TM3>
F:134-135/Region: DR motif
F:155-175/Domain: transmembrane #status predicted <TM4>
F:209-229/Domain: transmembrane #status predicted <TM5>
F:246-266/Domain: transmembrane #status predicted <TM6>
F:295-306/Domain: transmembrane #status predicted <TM7>
F:14,273/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:150,231/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predict
F:330/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 15.2%; Score 325; DB 2; Length 353;
Best Local Similarity 27.5%; Pred. No. 2.3e-18;
Matches 87; Conservative 59; Mismatches 110; Indels 60; Gaps 12;

QY 115 IYVLFVGVGVSNIYTLWKLRLRKSISLIVFHTNLAIADLFCVTLPPKIAHYLNGNNW 174

DB 44 LVYALFVLGIPGNAIVFMFGFKWKTVTTLWFLNLAIADLFCVTLPLIYISYVALSPHW 103

QY 175 VFGVMCRITTVVFGYNNYCAILLITCMGINRYLATAPFTYQKLPKRSF--SLLMCGIV 232

DB 104 PGRWLCKLNSIAOLNPFSSVFFLTVISLDRIYHLIHPGLSH--PHRTLKNKSLLVLFV 161

QY 233 WYMYFLYMLPFVILK-----OYHLVHSEITTCNDVVDACESPSSFRFY 277

DB 162 WLLASLLGGPTLYFRDTEVNNRIICNNFQYEL-----TLMRHV 203

QY 278 FVSLAP-FGLIPFVIIIFCYTTLIHLK-----SKRIWGLYIKAVLLILVITFCFA 330

DB 204 LTWVFLFGLYLLPLTMSSCYLCIFKTKQNILSSKHLWM-----ILSVVIAFMVCWT 258

QY 331 PTNII-----LVTHANYYYHNTDSLYFMYLIAICLGLNSCLDPPYLFVMSKVVDQLNFX 386

DB 259 PPHLSIWELSIHH-NSSFQNV--LQGGIPLSTGLAFLNSCLNPILYIISKKF-QARPR 314

QY 387 SAMARPLXRRPRDIWE 402

A;Accession: JH0622
A;Molecule type: DNA
A;Residues: 1-6, 'I', 8-165, 'H', 167-172, 'E', 174-204, 'V', 206-231, 'T', 233-238, 'F', 240-359 <E>
A;Cross-references: GB:S37491; NID:G249947; PIDN:AAB22270.1; PID:G249948
A;Experimental source: Balb/c
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F;30-53/Domain: transmembrane #status predicted <TM1>
F;65-90/Domain: transmembrane #status predicted <TM2>
F;103-123/Domain: transmembrane #status predicted <TM3>
F;147-166/Domain: transmembrane #status predicted <TM4>
F;195-218/Domain: transmembrane #status predicted <TM5>
F;240-263/Domain: transmembrane #status predicted <TM6>
F;278-300/Domain: transmembrane #status predicted <TM7>
F;4,176,188/Binding site: carboxylate (Asn) (covalent) #status predicted
F;141/Binding site: phosphate (Thr) (covalent) #status predicted
F;331,338,348/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 15.2%; Score 324; DB 2; Length 359;
Best Local Similarity 28.1%; Pred. No. 2.8e-18;
Matches 91; Conservative 52; Mismatches 125; Indels 56; Gaps 9;

Qy 87 EDGISLTHVNNATIGYLRSSLSQVTPAIYILLFVVGVPSN---IVTLWKLRLTKSIS 142
Db 8 EDGKRIQDDCPKAG--RNYIFVMPITLISYIFVGVFGNSLVIVIFYMKLKTA-- 63

Qy 143 LVIFHTNLAIADLLFCVTLPPFKIAYHLGNVNVFGEVMCRITTVFVGYNMYCAILITCM 202
Db 64 -SVFLNLALADLCFLTTLPLNAVYTA MEYQVPGNHLCKIASASVFNLYASVFLLTCL 122

Qy 203 GINRYIATAHPFTYQKLPKRSFLMCGIIVWVFLYMLPFVLKOEYHLVHSEITTCND 262
Db 123 SIDRYLAIVHPKMSRLRLTMLVAKVTCIIILMAGLASLPAVIYRNRYFIANTNITVC-- 180

Qy 263 VVDACESPSFRFY-----FVSLAPFGFLIPFVIIIFCVTTLIHLKLSKDRIW 311
Db 181 -----APHYSQNSTPIGUGLKNILGFEPFPIILTSYTLWKAUKAYKIQ 229

Qy 312 -----LQYIKAVLLILVIFITCPAPTNI-----ILVIHANYYYHNTDSLYFM 354
Db 230 KNIPRNDIGRI--INAILVLPFFSVVPHQIFSLDVLQLGVIHDCI----ADVVDTA 283

Qy 355 YLIALCGSLNSCLDPFLFYFVMSK 378
Db 284 MPITICIAVFNCLNPLFYGFGLCK 307

RESULT 32
I53033
G protein-coupled receptor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C;Accession: I53033
R;Heiber, M.; Docherty, J.M.; Shah, G.; Nguyen, T.; Cheng, R.; Heng, H.H.Q.; Marchese, F.
DNA Cell Biol. 14, 25-35, 1995
A;Title: Isolation of three novel human genes encoding G protein-coupled receptors.
A;Reference number: I53033; MUID:95134353; PMID:7832990
A;Accession: I53033
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-354 <RES>
A;Cross-references: GB:I36148; NID:G598152; PIDN:AAA63180.1; PID:G598153
C;Superfamily: G protein-coupled receptor 4
C;Keywords: G protein-coupled receptor

Query Match 15.1%; Score 321.5; DB 2; Length 354;
Best Local Similarity 27.6%; Pred. No. 4.4e-18;
Matches 79; Conservative 59; Mismatches 113; Indels 35; Gaps 6;

Qy 113 PAIYILLFVVGVPSNIVTLWKLRLTKSIS-LVIFHTNLAIADLLFCVTLPPFKIAYHLNG 171
Db 21 PSLYIFVIGVGLPTNCLALWAAYRQVRNELGVLMNLSTIADLLYICTLPLWVDYFLHH 80

Db 123 STDYLAIVHPMKSLRRRTMLVAKVTCIIIIWLAGLASLPALIIHRNVFFIENTNITVC-- 180
QY 263 VVDACESPSFRFY-----FVSLAFPGFLIPFVIIIFCVTTIIHLKSKDRITW 311
Db 181 -----AFHYESQNSTLPIGLGKTNILGFLFPFLIILTSYTLIWKALKRAYEIQ 229
QY 312 LG-----YIKAVLLILVITTCFAPTNIIL-----VIHANYYYHNTDSLYFMVL 356
Db 230 KMKPRNDDIFKIMAVLVEFFFSWPHQIFTELDVLQIQLGVHDCRI-----ADIVDTAMP 285
QY 357 IALCLGSLNSCLDPFLYFVMSK 378
Db 286 ITICIAFYFNCLNPLFYGLGK 307

RESULT 36
S15403
angiotensin II receptor type 1 - bovine
C:Species: Bos primigenius indicus x Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S15403
R:Sasaki, K.; Yamano, Y.; Bardhan, S.; Iwai, N.; Murray, J.J.; Hasegawa, M.; Matsuda, Y.
Nature 351, 230-233, 1991
A:Title: Cloning and expression of a complementary DNA encoding a bovine adrenal angiotensin II receptor type 1
A:Reference number: S15403; MUID:91251900; PMID:2041569
A:Accession: S15403
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-359 <SAS>
A:Cross-references: GB:X62294; NID:943; PIDN:CAA44182.1; PID:944
C:Superfamily: vertebrate rhodopsin

Query Match 15.0%; Score 320; DB 2; Length 359;
Best Local Similarity 27.6%; Pred. No. 5.9e-18;
Matches 85; Conservative 50; Mismatches 103; Indels 70; Gaps 8;

QY 111 VIPAIYILFVVGVPNS-----IVTLWKLRLTKSISLVIFHTNLATADLLFCVTLPPFKIA 166
Db 30 MIPTLYSIIFVVGIFGNSLWVIYFYMKLKTVA---SVFLNLALADCLFTLPLMAV 86
QY 167 YHLGNWVFGVWMCRTITVFGNNMYCAIILITCMGINRYLATAHPFTYQKLPKSPSL 226
Db 87 YTAEMVRWPGNYLCKIASASVSFNLYASVELLTCLSDRYLAIVHPMKSLRRRTMLVAK 146
QY 227 LMCGIIVWVFLYMLPFVILKQBYHLVHSEITTCDDWDACESPSFRFY----- 277
Db 147 VTICIIWLAGLASLPTIIHRNVFFIENTNITVC-----AFHYESQNSTLPV 193
QY 278 --FVSLAFPGFLIPFVIIIFCVTTIIHLKLS-----KDIRWLGVIKAVLLILVIF 325
Db 194 GLGLTKNIGLFLFPFLIILTSYTLIWKALKKAYEIQKPRKDDIF---KIILATVLPF 249
QY 326 TICFAPTNIILVIHANYYYHNTDSLYFMVL-----IALCLGSLNSCLDP 370
Db 250 FFSWPHQIFTEM-----DVLIIQLGRDKIEDIVDTAMPITICLAYFNCLNP 299
QY 371 FLYFVMSK 378
Db 300 LFYGLGK 307

RESULT 37
JC5653
G protein-coupled receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 17-Nov-1997 #sequence_revision 17-Nov-1997 #text_change 21-Jul-2000
C:Accession: JC5653
R:Ferrer-Martinez, A.; Felipe, A.; Mata, J.F.; Casado, F.J.; Pastor-Anglada, M.
Biochem. Biophys. Res. Commun. 238, 107-112, 1997
A:Title: Molecular cloning of a bovine renal G-protein coupled receptor gene (BRGR): Res
A:Reference number: JC5653; MUID:97445134; PMID:9299461
A:Accession: JC5653
A:Molecule type: mRNA

A:Residues: 1-361 <FER>
A:Cross-references: GB:U88366; NID:g2827875; PIDN:AA05611.1; PID:g2827876
A:Experimental source: renal epithelial cell
C:Superfamily: G protein-coupled receptor 4
C:Keywords: glycoprotein; phosphoprotein
F:24-46/Domain: transmembrane #status predicted <TM1>
F:58-79/Domain: transmembrane #status predicted <TM2>
F:96-117/Domain: transmembrane #status predicted <TM3>
F:137-153/Domain: transmembrane #status predicted <TM4>
F:189-207/Domain: transmembrane #status predicted <TM5>
F:229-253/Domain: transmembrane #status predicted <TM6>
F:265-289/Domain: transmembrane #status predicted <TM7>
F:3,8,62/Binding site: carbonylrate (Asn) (covalent) #status predicted
F:134,218,293,313/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status pre
F:204,221,328/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pre

Query Match 15.0%; Score 319.5; DB 2; Length 361;
Best Local Similarity 26.7%; Pred. No. 6.5e-18;
Matches 76; Conservative 65; Mismatches 107; Indels 37; Gaps 8;

QY 113 PAIYILLFVVGVPNSIVTLWKLRLTKSIS-LVIFHTNLATADLLFCVTLPPFKIAHYLNG 171
Db 25 PVVYVWLVVGFNPANCLSLYYGYLQIKARNELGVVLCNLTVADLFYICSLPFWLQVLOH 84
QY 172 NNWVGEVWMCRTITVFGNNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSLMCGI 231
Db 85 DHWSHDDLSCQVCGILLYENIYISVGLCCISIDRYLAVAHFPHQFRTLKAAMGVSA 144
QY 232 VWMVFLYMLPFVILKQBYHLVHSEITTCDDWDAC-----ESPSSFRFYFYVSLAFF 284
Db 145 IWKELLTSI-----YFLMHSEVVEDADRHRVCFEHYPLEPRQGINIYRFL---V 192
QY 285 GFLIPFVIIIFCVTTIIHLK-----SKDIRWLGVIKAVLLILVITTCFAPTNIIL 336
Db 193 GFLPFCILLASRYGILRAVRSHCTQSKRDKQIO---RLVLSIVLFLACFLPYHVL 248
QY 337 VIHANYYYHNTD---SLYFMYLIACLGSLNSCLDPFLYFVMSK 378
Db 249 LVR--SLWESSCDFAKGIFNAYHFSLTLTSFNCAVDPLVLCFVSE 291

RESULT 38
S68208
G protein-coupled receptor 12A - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C:Accession: S68208
R:An, S.; Tsai, C.; Goetzl, E.J.
FEBS Lett. 375, 121-124, 1995
A:Title: Cloning, sequencing and tissue distribution of two related G protein-coupled re
A:Reference number: S68207; MUID:96087098; PMID:7498459
A:Accession: S68208
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-365 <ANS>
A:Cross-references: EMBL:U35398; NID:G1015418; PIDN:AAA79060.1; PID:G1015419
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

Query Match 14.9%; Score 318.5; DB 2; Length 365;
Best Local Similarity 26.8%; Pred. No. 7.9e-18;
Matches 77; Conservative 67; Mismatches 102; Indels 41; Gaps 10;

QY 113 PAIYILLFVVGVPNSIVTLWKLRLTKSIS-LVIFHTNLATADLLFCVTLPPFKIAHYLNG 171
Db 25 PVVYVWLVVGFNPANCLSLYYGYLQIKARNELGVVLCNLTVADLFYICSLPFWLQVLOH 84
QY 172 NNWVGEVWMCRTITVFGNNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSLMCGI 231
Db 85 DNWSHDDLSCQVCGILLYENIYISVGLCCISIDRYLAVAHFPHQFRTLKAAMGVSA 144
QY 232 VWMVFLYMLPFVILKQBYHLVHSEITTCDDWDAC-----TTC--HDVVDAACESPSFRFYVSLA 282

Db 145 IWAKELLTST-----YFLMHEEVEDENQHRVCFEHPYIQAWO--RAINYRFL--- 191
QY 283 FGFLLPFFVLIIFCYTTLIHKL-----SKDRWLGVYKAVLLILVIFTCAPTNI 334
Db 192 -VGFLLPCLLLASYOGILRAVRSHGTSKRDKIQ---RLVLTSTVFIACFLPFLYHV 246
QY 335 ILVIHHYYHNTD---SLYFMYLIALCIGLSNCLDPLFLFYVMSK 378
Db 247 LLLVR--SVWEASCDPAKGVFNAYHESLLLTSCNVADPVLICVSE 291

RESULT 39

A53752
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: A53752
R:Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarro J. Biol. Chem. 269, 12391-12394, 1994
A:Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A:Reference number: A53752; MUID:94230294; PMID:8175642
A:Accession: A53752
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-358 <PRA>
A:Cross-references: GB:L24445; NID:g437661; PIDN:AAA31378.1; PID:g437662
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 14.9%; Score 318; DB 2; Length 358;
Best Local Similarity 26.7%; Pred. No. 8.4e-18;
Matches 92; Conservative 74; Mismatches 116; Indels 62; Gaps 15;
QY 62 TFEPLSLDIEGTGAT---TTI-KAECPEISITLHNNATIGVLRSSLSTQVIPAII 116
Db 10 SYEDF-FGDFSNYSYSTDPPTLLDSAPCRSESLT-----NSYVLIITY 53
QY 117 ILFLVGVPSN-IVTLWKLRLTKSISLVIFHTNLAIALDLFCVTLPEKIAVHLGNVW 175
Db 54 ILVFLSLGNSLVMLVILYSRSTCVTDVYLNLAIALDLFATLPIWAASKVHG--WT 111
QY 176 FGEVMCRTITVVFYGNMYCAILITCMGINRYLATAHPTFYQKLPKRSFLMCGIVWM 235
Db 112 FGTPLKWSLVKEVNFYSGILLACISVDRLAIVHA-TRTWIQKRLVKFICLSMVG 170
QY 236 VFLLMPLFVLKQYHLVHSEITTCDDVDACESPSSFRFYFVSLAFGLPFIPIIF 295
Db 171 SLILSLPILFLRNAIFPPNSS-PVCYE--DMGNSTAKRMVLRILPQTGFILPLVMLF 227
QY 296 CYTTLIHKLSKDRILWGY---IKAVLLIIVFTICFAPTNIILVIHHANYHHNTDSL 351
Db 228 CY---VFTLTFLQAHGQKHRAMRVIFAVLLFLCWLFPYNLVLD-----TDIL 274
QY 352 YFMYLI-----AL-----CLGSLNSCLDPLFLFYVMSK 378
Db 275 MRTHVIQETCERNRIDRALDATEILGLHSLCLNPIIYAFIGQK 318

RESULT 40

JC4800
P2Y6 receptor - human
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: JC4800; G02514
R:Communi, D.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 222, 303-308, 1996
A:Title: Cloning, functional expression and tissue distribution of the human P2Y6 receptor
A:Reference number: JC4800; MUID:96222498; PMID:8670200
A:Accession: JC4800
A:Molecule type: mRNA
A:Residues: 1-328 <COM>
A:Cross-references: EMBL:X97058
A:Experimental source: placenta

R:Hammet, F.; Southey, M.C.; Somers, G.R.; Hutchins, A.M.; Venter, D.J.
submitted to the EMBL Data Library, March 1996
A:Reference number: H01373
A:Accession: G02514
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 'M', 4-328 <HAM>
A:Cross-references: EMBL:U52464; NID:g1407632; PIDN:AAB03572.1; PID:g1407633
C:Genetics:
A:Gene: P2Y6
C:Superfamily: ATP receptor P2u
C:Keywords: glycoprotein; placenta; receptor; transmembrane protein
F:26-52/Domain: transmembrane #status predicted <TM1>
F:63-86/Domain: transmembrane #status predicted <TM2>
F:104-122/Domain: transmembrane #status predicted <TM3>
F:143-167/Domain: transmembrane #status predicted <TM4>
F:193-216/Domain: transmembrane #status predicted <TM5>
F:241-264/Domain: transmembrane #status predicted <TM6>
F:283-305/Domain: transmembrane #status predicted <TM7>
F:5,173/Binding site: carbonylrate (Asn) (covalent) #status predicted

Query Match 14.9%; Score 317.5; DB 2; Length 328;
Best Local Similarity 27.7%; Pred. No. 8.4e-18;
Matches 83; Conservative 53; Mismatches 125; Indels 39; Gaps 9;
QY 104 RSSLSSTQVIPAIIYILLFVGVPSNIVTLWKLRLTKSISLVIFHT-NLAIALDLFCVTL 162
Db 21 RENFKQLLPPVYSAVLAAGPLNICVITQTSRRALTRAVYTLNLAIALDLFACSLP 80
QY 163 FKIAVHLGNVWVGEVNMCRITTVVFYGNMYCAILITCMGINRYLATAHPT-YOKLPK 221
Db 81 LLINYNAQGDHWPFGDFACRLVRLFYANLHGSILFLTCISFORVILGICHPLAPWHKRG 140
QY 222 RSFSLMCGIVWMVFLYMLPVLKQYHLVHSEITTCDDVDACESPSSFRFYF--- 278
Db 141 RRAAWLVCVAVMLAVTTQCLPTAIPAATG--IQNRNIVCYDL-----SPPALATHYMPYG 193
QY 279 VSLAFPGFLIPFVLIIFCYTTLIHKLSKDRILWGYIKAV-----LLILVIFT 326
Db 194 MALTVIGFLFLPFAALLACYCLLACLKQD---GPAEPVAQERKGAARMVAVVAAFA 249
QY 327 ICFAPTNIILVIHHANYHHNTDSLFLYFMYLIALC-----LGSLSNCLDPLFLFYVMSK 378
Db 250 ISFLPFHITKTAYLA---VRSTPGVPCVLEAFAAAYKGRPFASANSVLDPIFLFYFTQK 306

Search completed: June 24, 2003, 12:06:34
Job time : 28.9681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:05:46 ; Search time 32.4601 Seconds
(without alignments)
1356.747 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136

Sequence: 1 TLTYXQHPVAGSQDKRMKIL.....AMARPLXPRRDIWEDIHAW 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 10826813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	509	23.8	425	10	US-09-782-980-80
2	509	23.8	425	10	US-09-884-430-4
3	493	23.1	359	9	US-10-094-417-10
4	493	23.1	359	9	US-09-782-974C-76
5	493	23.1	359	10	US-09-739-151-2
6	489	22.9	359	9	US-10-190-469-1
7	444.5	20.8	359	10	US-09-943-718-6
8	431.5	20.2	361	9	US-10-222-024-2
9	431.5	20.2	361	9	US-10-251-385-78
10	428.5	20.1	361	9	US-10-251-385-206
11	426.5	20.0	348	10	US-09-827-937A-17
12	399	18.7	362	9	US-10-092-135-3
13	399	18.7	362	9	US-09-779-679-28
14	398.5	18.7	367	9	US-09-828-478-6
15	395	18.5	339	9	US-09-828-478-4
16	395	18.5	339	9	US-10-251-385-32
17	395	18.5	339	10	US-09-848-889-12
18	395	18.5	339	10	US-09-788-133-2
19	395	18.5	362	9	US-10-092-135-4

20	393	18.4	537	9	US-10-311-956-4	Sequence 4, Appli
21	389	18.2	339	9	US-10-251-385-182	Sequence 182, App
22	388	18.2	366	9	US-09-779-679-25	Sequence 25, Appl
23	378.5	17.7	330	10	US-09-826-791-2	Sequence 2, Appli
24	378.5	17.7	346	9	US-09-828-478-2	Sequence 2, Appli
25	378.5	17.7	346	9	US-09-779-679-2	Sequence 2, Appli
26	378.5	17.7	346	9	US-09-779-679-26	Sequence 26, Appl
27	378.5	17.7	346	10	US-09-826-791-6	Sequence 6, Appli
28	378.5	17.7	346	10	US-09-866-230-7	Sequence 7, Appli
29	378.5	17.7	346	10	US-09-866-230-9	Sequence 9, Appli
30	378	17.7	337	9	US-09-828-478-5	Sequence 5, Appli
31	378	17.7	337	10	US-09-866-230-8	Sequence 8, Appli
32	377.5	17.7	373	9	US-10-092-135-7	Sequence 7, Appli
33	377	17.6	302	9	US-10-024-494-30	Sequence 30, Appl
34	377	17.6	337	9	US-09-779-679-27	Sequence 27, Appl
35	376.5	17.6	373	9	US-10-092-135-6	Sequence 6, Appli
36	372	17.4	373	9	US-10-092-135-5	Sequence 5, Appli
37	371.5	17.4	360	10	US-09-943-718-4	Sequence 4, Appli
38	370	17.3	337	9	US-10-167-192-3	Sequence 3, Appli
39	369.5	17.3	337	9	US-10-023-775B-2	Sequence 2, Appli
40	369.5	17.3	337	9	US-10-270-144-2	Sequence 2, Appli
41	369.5	17.3	337	9	US-10-188-405-8	Sequence 8, Appli
42	369.5	17.3	337	9	US-09-885-453-1	Sequence 1, Appli
43	369.5	17.3	337	9	US-10-079-384-14	Sequence 14, Appl
44	369.5	17.3	337	10	US-09-943-798-4	Sequence 4, Appli
45	369.5	17.3	341	9	US-10-270-587-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-782-980-80
; Sequence 80, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STMTST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121

;; PRIOR FILING DATE: 1998-05-29
;; PRIOR APPLICATION NUMBER: 09/672,721
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: 09/049,799
;; PRIOR FILING DATE: 1998-03-27
;; NUMBER OF SEQ ID NOS: 176
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 80
;; LENGTH: 425
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-782-380-80

Query Match 23.8%; Score 509; DB 10; Length 425;
Best Local Similarity 31.2%; Pred. No. 2e-36;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTPEEPFLSDI 71
DB 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRPNPNKYEPPWEDEE 60
QY 72 EGWTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNS 127
DB 61 KNESGLTEYRLVSNKSSPLQKLPAPFISEDASGYLTSSWLTFLFVPSVYTGTFVWSLPLN 120
QY 128 I--VTLWKLRLTKSISLVIFHTNLATADLLFCVTLPFKIAYHLNGNNVFGVEMCRIT 185
DB 121 IMAIVFILKMKVK-K-PAVVYMLHLATADVLVSVLPFKISYFSGSDWQFSGELCRFVT 179
QY 186 VVFYGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLLMCGIVVMVFLYMLP 242
DB 180 AAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
QY 243 FVILKQEHVHLVHSEITTCCHVDVADACESPSFRFYFVSLAFFGFLIPFVIIIFCYTTLIH 302
DB 237 LVLKEQTIQVPLNITTCCHVDVADACESPSFRFYFVSLAFFGFLIPFVIIIFCYTTLIH 295
QY 303 KLKS-----KDRILWGLYKAVLL---ILVIFTCFAPTNIILVIHANYYYH---NTD 349
DB 296 CLSSAVANRSKSR-----ALFLSAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
DB 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 2
US-09-884-430-4
;; Sequence 4, Application US/09884430
;; Patent No. US20020151046A1
;; GENERAL INFORMATION:
;; APPLICANT: Glucksmann, Maria Alexandra
;; TITLE OF INVENTION: SANTIAGO, IMMACULADA SILOS
;; FILE OF INVENTION: 52871, A NOVEL HUMAN G PROTEIN COUPLED
;; FILE REFERENCE: MNT-165
;; CURRENT APPLICATION NUMBER: US/09/884,430
;; PRIOR FILING DATE: 2001-06-18
;; PRIOR APPLICATION NUMBER: USSN 60/212,331
;; PRIOR FILING DATE: 2000-06-16
;; PRIOR APPLICATION NUMBER: USSN 60/269,758
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 425
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-884-430-4

Query Match 23.8%; Score 509; DB 10; Length 425;
Best Local Similarity 31.2%; Pred. No. 2e-36;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTPEEPFLSDI 71
DB 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRPNPNKYEPPWEDEE 60
QY 72 EGWTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNS 127
DB 61 KNESGLTEYRLVSNKSSPLQKLPAPFISEDASGYLTSSWLTFLFVPSVYTGTFVWSLPLN 120
QY 128 I--VTLWKLRLTKSISLVIFHTNLATADLLFCVTLPFKIAYHLNGNNVFGVEMCRIT 185
DB 121 IMAIVFILKMKVK-K-PAVVYMLHLATADVLVSVLPFKISYFSGSDWQFSGELCRFVT 179
QY 186 VVFYGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLLMCGIVVMVFLYMLP 242
DB 180 AAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
QY 243 FVILKQEHVHLVHSEITTCCHVDVADACESPSFRFYFVSLAFFGFLIPFVIIIFCYTTLIH 302
DB 237 LVLKEQTIQVPLNITTCCHVDVADACESPSFRFYFVSLAFFGFLIPFVIIIFCYTTLIH 295
QY 303 KLKS-----KDRILWGLYKAVLL---ILVIFTCFAPTNIILVIHANYYYH---NTD 349
DB 296 CLSSAVANRSKSR-----ALFLSAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
DB 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 3
US-10-094-417-10
;; Sequence 10, Application US/10094417
;; Publication No. US20030045685A1
;; GENERAL INFORMATION:
;; APPLICANT: Tian, Hui
;; APPLICANT: Zhao, Jiagang
;; APPLICANT: Chen, Jin-Long
;; APPLICANT: Cutler, Gene
;; APPLICANT: Tularik Inc.
;; TITLE OF INVENTION: NO. US20030045685A1el Receptors
;; FILE REFERENCE: 018781-008110US
;; CURRENT APPLICATION NUMBER: US/10/094,417
;; CURRENT FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: US 09/802,803
;; PRIOR FILING DATE: 2001-03-09
;; PRIOR APPLICATION NUMBER: US 60/276,649
;; PRIOR FILING DATE: 2001-03-16
;; NUMBER OF SEQ ID NOS: 43
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 10
;; LENGTH: 359
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
US-10-094-417-10

Query Match 23.1%; Score 493; DB 9; Length 359;
Best Local Similarity 35.7%; Pred. No. 4.1e-35;
Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;
QY 96 NNATIGYLRSSLSSTQVIPAIYILLFVVGVPNSITVTLWKLRLTKSIS-LVIFHTNLAIAD 154
DB 10 DNATLQMLRNPAFAVALPVVYSLVAAVSIPIGNLFLSLVLCRRMGPRSPSVIFMINLSVTD 69
QY 155 LFLCVTLTPFKIAYHLNGNNVFGVEMCRITTVVYGNMYCAIILTCMGINRYLATAHP 214
DB 70 LMLASVLPFOIYYHCNRRHHWFGVLLCNVTVVAFYANMYSSILLTMTICISVERELGVLYPL 129
QY 215 TYQKLPKRSFSLLMCGIVVMVFLYMLPFLVILKQEHVHLVHSEITTCCHVDVADACESPS--- 271
DB 130 SSKRRRRRYAAACAGTWTLLLTALSPILARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189

[illegible]

```

RESULT 8
US-10-222-024-2
; Sequence 2, Application US/10222024
; Publication No. US20030104487A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (GB)
; APPLICANT: Pfizer Ltd. (GB)
; TITLE OF INVENTION: EP except GB, US, JP)
; FILE REFERENCE: PCS22032
; CURRENT APPLICATION NUMBER: US/10/222,024
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: GB 0119920.7
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-024-2

```

Query Match	.20.2%	Score 431.5;	DB 9;	Length 361;
Best Local Similarity	31.6%;	Pred.No. 9.5e-30;		
Matches	95;	Conservative 64;	Mismatches 107;	Indels 35; Gaps 9;
Qy	102	YLRSSLSGVIPALYITLLFVVGVSINVTLWKLSLRTKSI-SLVIPHTNLAIADLLFCVT	160	
Db	24	YAHSHTRIVMPLHYSLVFIIGLVGNLLALVVIVQRKKINSTLTSTNVLISDILFTTA	83	
Qy	161	LPPKIAVHLGNXNWVGEVMCRITTVFYGYMYCAIILTCMGINRYLATATAHPPTYOKLP	220	
Db	84	LPTRIAAYAMGFOWRIGDALCRITALVFYINTVAGNVFMTCISIDREFAVVVPLRYNNKI	143	
Qy	221	KRFSSLMLCGIVMMVPLYMLPFVI---LKOEYHLVHSEITTHCHDVVDACESSSFRFY	277	
Db	144	RIBHAKGVCIFWMILVFAQYQLPLINPMSKOE-----AERITCMENPFBETKS--LPWI	196	
Qy	278	FVSLAFEGFLIPFVIIIFCYTTLIHK-----LKSKDRIMLGVIKAVLLILVIFTIC	328	
Db	197	LLGACFGYVYLPIIIICYSQICKLFRTAQKNPLTEKSGVNKALNTIILIVFVLC	256	
Qy	329	FAPTNITLVTH-----HANY-----YYHNTD-SLYFWNYIALCLGSUNSCLDPLPYFMS	377	
Db	257	FTYPVVAIOHMIIKKLFNSFLBCSRHSFQISLHF-----TVCMLMNFCMCDPEIFYFAC	312	

Qy 378 K 378
Db 313 K 313

RESULT 9

US-10-251-385-78
; Sequence 78, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-78

Query Match 20.2%; Score 431.5; DB 9; Length 361;
Best Local Similarity 31.6%; Pred. No. 9.5e-30;
Matches 95; Conservative 64; Mismatches 107; Indels 35; Gaps 9;

Qy 102 YLRSSLSTQVPAIYILLFVVGPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
Db 24 YAHSTARIWPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVSDILFTTA 83
Qy 161 LPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAIILTCMGINRYLATAHPTFYOKLP 220
Db 84 LPTRIAAYAMGDFWRIGDLCRITALVFYINTYAGVNFMTCLSIDRFAIVVHPLRYNKK 143
Qy 221 KRFSLLMCGIWMVFLYMLPFVI---LKQEYHLVHSEITTCDDVVDACESPSSFRFY 277
Db 144 RIEHAKGVCIFWILVFAQTLLPLINPMKOE-----AERITCMEYFNFEETKS--LPWI 196
Qy 278 FVSLAFFGFLIPFVILIFCYTTLIHLK-----LKSODRIWLGYIKAVLLIVITIC 328
Db 197 LIGACFIGVVLPLIILICYSQICCKLFRKAKQNPTEKSGVKKAKN---TVCLMNFNCMD 256
Qy 329 FAPTNIIIVIH-----HANY-----YYHNTD-SLYFMYLIALCLGSLNSCLDPLFYFMS 377
Db 257 FTPYHVAIQHMVKLRFSNFLECSQRHSFQISLHF---TVCLMNFNCMDPIIYFFAC 312
Qy 378 K 378
Db 313 K 313

RESULT 10

US-10-251-385-206
; Sequence 206, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496

; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-206

Query Match 20.1%; Score 428.5; DB 9; Length 361;
Best Local Similarity 31.1%; Pred. No. 1.7e-29;
Matches 96; Conservative 64; Mismatches 98; Indels 51; Gaps 10;
Qy 102 YLRSSLSTQVPAIYILLFVVGPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
Db 24 YAHSTARIWPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVSDILFTTA 83
Qy 161 LPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAIILTCMGINRYLATAHPTFYOKLP 220
Db 84 LPTRIAAYAMGDFWRIGDLCRITALVFYINTYAGVNFMTCLSIDRFAIVVHPLRYNKK 143
Qy 221 KRFSLLMCGIWMVFLYMLPFVI---LKQEYHLVHSEITTCDDVVDACESPSSFRFY 277
Db 144 RIEHAKGVCIFWILVFAQTLLPLINPMKOE-----AERITCMEYFNFEETKS--LPWI 196
Qy 278 FVSLAFFGFLIPFVILIFCYTTLIHLK-----LKSODRIWLGYIKAVLL 320
Db 197 LIGACFIGVVLPLIILICYSQICCKLFRKAKQNPTEKSGVKKAKN---TVCLMNFNCMD 248
Qy 321 ILVIFTCFAPTNIIIVIH-----HANY-----YYHNTD-SLYFMYLIALCLGSLNSCLD 369
Db 249 IIVVFLCTPYHVAIQHMVKLRFSNFLECSQRHSFQISLHF---TVCLMNFNCMD 304
Qy 370 PLYFYVMSK 378
Db 305 PFYIYFFACK 313

RESULT 11

US-09-827-937A-17
; Sequence 17, Application US/09827937A
; Patent No. US20020052043A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human G-Protein Coupled Receptors
; FILE REFERENCE: 1488.1220003
; CURRENT APPLICATION NUMBER: US/09/827,937A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 08/852,824
; PRIOR FILING DATE: 1997-05-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-937A-17

Query Match 20.0%; Score 426.5; DB 10; Length 348;
Best Local Similarity 31.6%; Pred. No. 2.5e-29;
Matches 95; Conservative 63; Mismatches 108; Indels 35; Gaps 9;
Qy 102 YLRSSLSTQVPAIYILLFVVGPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
Db 21 YAHSTARIWPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVSDILFTTA 80
Qy 161 LPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAIILTCMGINRYLATAHPTFYOKLP 220
Db 81 LPTRIAAYAMGDFWRIGDLCRITALVFYINTYAGVNFMTCLSIDRFAIVVHPLRYNKK 140
Qy 221 KRFSLLMCGIWMVFLYMLPFVI---LKQEYHLVHSEITTCDDVVDACESPSSFRFY 277

```
Db 141 RIEHAKGVCIFWVILVFAQTLLPLLNPMNSKQE-----AERITCMEVFNFEETKS--LPWI 193
Qy 278 FVSIAFGFLPFFVILIFCVTTLLHK-----LKSDDRIGWGYKAVLLILVIFTIC 328
Db 194 LLGACFYGVLPLIIKICYSQICCKLFRKAKONPLTKSGVKNKALNTIILIVVFLC 253
Qy 329 FAPNTIILVIH-----HANY-----YVHTND-SLYFMYLIALCIGLSNCLDPFLYFVMS 377
Db 254 FFPVHVAIQHMIKKLRFSNFLECSQRHSQISLHF-----TVCLMNFNCMDPFIYFFAC 309
Qy 378 K 378
Db 310 K 310

RESULT 12
US-10-092-135-3
; Sequence 3, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
; TITLE OF INVENTION: HGPBMB27
; FILE REFERENCE: D0134.NP
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US/10/092,135
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/278,983
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 362
; TYPE: PRT
; ORGANISM: GALLUS GALLUS
US-10-092-135-3

Query Match 18.7%; Score 399; DB 9; Length 362;
Best Local Similarity 28.3%; Pred. No. 6.5e-27;
Matches 91; Conservative 63; Mismatches 122; Indels 46; Gaps 10;

Qy 73 GWTGATTITKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIYIILLFVGVPSNIVTLW 132
Db 21 GWAAGNATTKC-----SLTKTGQFYVPLTVYILVFTIGFLGNSVAIW 63
Qy 133 KLSLRK---SISLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNNWVGEVMSRITTVVFY 189
Db 64 MFVFMHRPWSGISVYMF--NLALADFLYVLTLPALIFYFNKTDWIFGDMCKLQRFIFH 121
Qy 190 GNMVCAILILTCMGINRYLATAHPTVQKLPKRSFSLLMCGIWMVVMVFLYMLPFVILKQE 249
Db 122 VNLGYSILFLTCISVHRYTGVVHPLKSLGRKKKNVYVSSLVWALVAVIAP-ILFYSG 180
Qy 250 YHLVHSEITTCDDVDACESPSSFRFYFVSLAFFGLIPFVIIIFCVYTLIHLKSKD- 308
Db 181 TGVRRKNTICYDTT-ADEYLSY-FVYSMCTTVFMFCIPFIVILGCGYGLIVKALYKDL 238
Qy 309 -----RIWLGYIKAVLLIIVFTICFAPTNIIILVIH-HANYYYHN-----TDSLYF 353
Db 239 DNSPLRRKSIYL-----VIIIVTFVAVSYLPFHVMKTLNLRARLDFQTPQMCAFNKVKYA 293
Qy 354 MYLIALCIGLSNCLDPFLYFV 375
Db 294 TYQVTRGLASLNSCVDPILYFL 315

RESULT 13
US-09-779-679-28
; Sequence 28, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
```

```
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-779-679-28

Query Match 18.7%; Score 399; DB 9; Length 362;
Best Local Similarity 28.3%; Pred. No. 6.5e-27;
Matches 91; Conservative 63; Mismatches 122; Indels 46; Gaps 10;

Qy 73 GWTGATTITKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIYIILLFVGVPSNIVTLW 132
Db 21 GWAAGNATTKC-----SLTKTGQFYVPLTVYILVFTIGFLGNSVAIW 63
Qy 133 KLSLRK---SISLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNNWVGEVMSRITTVVFY 189
Db 64 MFVFMHRPWSGISVYMF--NLALADFLYVLTLPALIFYFNKTDWIFGDMCKLQRFIFH 121
Qy 190 GNMVCAILILTCMGINRYLATAHPTVQKLPKRSFSLLMCGIWMVVMVFLYMLPFVILKQE 249
Db 122 VNLGYSILFLTCISVHRYTGVVHPLKSLGRKKKNVYVSSLVWALVAVIAP-ILFYSG 180
Qy 250 YHLVHSEITTCDDVDACESPSSFRFYFVSLAFFGLIPFVIIIFCVYTLIHLKSKD- 308
Db 181 TGVRRKNTICYDTT-ADEYLSY-FVYSMCTTVFMFCIPFIVILGCGYGLIVKALYKDL 238
Qy 309 -----RIWLGYIKAVLLIIVFTICFAPTNIIILVIH-HANYYYHN-----TDSLYF 353
Db 239 DNSPLRRKSIYL-----VIIIVTFVAVSYLPFHVMKTLNLRARLDFQTPQMCAFNKVKYA 293
Qy 354 MYLIALCIGLSNCLDPFLYFV 375
Db 294 TYQVTRGLASLNSCVDPILYFL 315

RESULT 14
```

```
US-09-828-478-6
; Sequence 6, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match      18.7%; Score 398.5; DB 9; Length 367;
Best Local Similarity 28.3%; Pred. No. 7.3e-27;
Matches 99; Conservative 60; Mismatches 136; Indels 55; Gaps 9;

QY 42 SDNSAKPTLTIKSFNG---GPQNTFEFFPLSDIEGWTGATTIKAECPEDSISTLHVNNA 98
Db 23 SDSS-----QSMNGLEVAPPGLITNFSLATAE-----QCQGE-----54
QY 99 TIGYLRSSLSQTQVIPAIIYLLFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFF 157
Db 55 -----TPLENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSC 108
QY 158 CVTLPPFKIAVHLNGNWMVFGVMCRITTVVFGYGNMYCAIILITCMGINRYLATAPHTYQ 217
Db 109 VLVLTPLRVVHFGSNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSL 168
QY 218 KLPKRSPSLMCGTVMVWVFLYMLPPFVILKQYHLVHSEITTCCHDVVDACESPSPFRFY 277
Db 169 KLRPLVAHLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHA 221
QY 278 FVSLAPFGFLIPVFIIFCYTTLIHKLKS-----KDRWLGVYKAVLLILVIFTCFAPT 333
Db 222 LVSLA-VAFTFPFITTVTCYLLIIRSLRQGLRVEKRLTKAVRMIAIVLAIFLVCFVPYH 280
QY 334 IILVHIANYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 281 VNRSVYVHLVHSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 330

RESULT 15
US-09-828-478-4
; Sequence 4, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-4

Query Match      18.5%; Score 395; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

QY 107 LSTQVIPAIYILLFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFCVTLPFKI 165
Db 29 LENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSCVLLVPLTRL 88
QY 166 AYHLNGNWMVFGVMCRITTVVFGYGNMYCAIILITCMGINRYLATAPHTYQKLPKRSPS 225
Db 89 VVHFGSNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSLKLRPLYA 148
QY 226 LLMCGIWMVWVFLYMLPPFVILKQYHLVHSEITTCCHDVVDACESPSPFRFYVSLA 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHALVSLA-VA 200
QY 286 FLIPFVIIIFCYTTLIHKLKS-----KDRWLGVYKAVLLILVIFTCFAPTNIILVIHHA 341
Db 201 FTFPFITTVTCYLLIIRSLRQGLRVEKRLTKAVRMIAIVLAIFLVCFVPYHVNRSVYVL 260
QY 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 261 HVRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 302

RESULT 16
US-10-251-385-32
; Sequence 32, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-32

Query Match      18.5%; Score 395; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

QY 107 LSTQVIPAIYILLFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFCVTLPFKI 165
Db 29 LENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSCVLLVPLTRL 88
QY 166 AYHLNGNWMVFGVMCRITTVVFGYGNMYCAIILITCMGINRYLATAPHTYQKLPKRSPS 225
Db 89 VVHFGSNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSLKLRPLYA 148
QY 226 LLMCGIWMVWVFLYMLPPFVILKQYHLVHSEITTCCHDVVDACESPSPFRFYVSLA 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHALVSLA-VA 200
QY 286 FLIPFVIIIFCYTTLIHKLKS-----KDRWLGVYKAVLLILVIFTCFAPTNIILVIHHA 341
Db 201 FTFPFITTVTCYLLIIRSLRQGLRVEKRLTKAVRMIAIVLAIFLVCFVPYHVNRSVYVL 260
QY 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 261 HVRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 302
```

```
US-09-828-478-6
; Sequence 6, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match      18.7%; Score 398.5; DB 9; Length 367;
Best Local Similarity 28.3%; Pred. No. 7.3e-27;
Matches 99; Conservative 60; Mismatches 136; Indels 55; Gaps 9;

QY 42 SDNSAKPTLTIKSFNG---GPQNTFEFFPLSDIEGWTGATTIKAECPEDSISTLHVNNA 98
Db 23 SDSS-----QSMNGLEVAPPGLITNFSLATAE-----QCQGE-----54
QY 99 TIGYLRSSLSQTQVIPAIIYLLFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFF 157
Db 55 -----TPLENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSC 108
QY 158 CVTLPPFKIAVHLNGNWMVFGVMCRITTVVFGYGNMYCAIILITCMGINRYLATAPHTYQ 217
Db 109 VLVLTPLRVVHFGSNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSL 168
QY 218 KLPKRSPSLMCGTVMVWVFLYMLPPFVILKQYHLVHSEITTCCHDVVDACESPSPFRFY 277
Db 169 KLRPLVAHLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHA 221
QY 278 FVSLAPFGFLIPVFIIFCYTTLIHKLKS-----KDRWLGVYKAVLLILVIFTCFAPT 333
Db 222 LVSLA-VAFTFPFITTVTCYLLIIRSLRQGLRVEKRLTKAVRMIAIVLAIFLVCFVPYH 280
QY 334 IILVHIANYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 281 VNRSVYVHLVHSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 330

RESULT 15
US-09-828-478-4
; Sequence 4, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-4

Query Match      18.5%; Score 395; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

QY 107 LSTQVIPAIYILLFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFCVTLPFKI 165
Db 29 LENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSCVLLVPLTRL 88
QY 166 AYHLNGNWMVFGVMCRITTVVFGYGNMYCAIILITCMGINRYLATAPHTYQKLPKRSPS 225
Db 89 VVHFGSNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSLKLRPLYA 148
QY 226 LLMCGIWMVWVFLYMLPPFVILKQYHLVHSEITTCCHDVVDACESPSPFRFYVSLA 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHALVSLA-VA 200
QY 286 FLIPFVIIIFCYTTLIHKLKS-----KDRWLGVYKAVLLILVIFTCFAPTNIILVIHHA 341
Db 201 FTFPFITTVTCYLLIIRSLRQGLRVEKRLTKAVRMIAIVLAIFLVCFVPYHVNRSVYVL 260
QY 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 261 HVRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 302
```

RESULT 17

US-09-848-889-12
; Sequence 12, Application US/09848889
; Patent No. US2002002555A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Cheng, Muzong
; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
; FILE REFERENCE: PC-0042 CIP
; CURRENT APPLICATION NUMBER: US/09/848,889
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2002002555A1 9992700
US-09-848-889-12

Query Match 18.5%; Score 395; DB 10; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

Qy 107 LSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRKTS-ISLVIFHTNLAIADLLFCVTLPLFKI 165
Db 29 LENMLFASFYLLDFILALVGNLTALWLFIRDHKSQTPANVFLMHLAVADLSCLVLPTRL 88

Qy 166 AYHLNGNNWVFGVCMCRITTVFVGNMYCAIILITCMGINRYLATATAHPFTYQKLPKRSFS 225
Db 89 VYHFGNHNWPFGEIACRLTGLFGLYNNMYSIYFLTCISADRFALVHPVKSLLRRPLA 148

Qy 226 LLMCGIYVMVFLYMLPFVILKQYHLVHSEITTCCHDVVDACESPSSFRFYFVSLAFFG 285
Db 149 HLACAFLLWVVAVAMAPLLVSPQ-----TVQTNHTVCLQLYREKASHHALVSLA-VA 200

Qy 286 FLIPFVIIIFCYTTLIHKLK-----KORIWLGYIKAVLLILVIFTICFAPTNIILVIHHA 341
Db 201 FTFFPITVTCLLIIRSLRQGLRVEKRLKTKAVRMIAIIVLFLVCFVYVHNRSVYVL 260

Qy 342 NYHYHTDS-----LYFWYLIALCLGSLNSCLDPLFYVMSK 378
Db 261 HYRSHGASCATORILALANRITSCLTSLNGALDPIMYFFVAE 302

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2002002555A1 9992700
US-09-848-889-12

RESULT 18

US-09-788-133-2
; Sequence 2, Application US/09788133
; Patent No. US20020052001A1
; GENERAL INFORMATION:
; APPLICANT: GLAXO GROUP LTD
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: P79011
; CURRENT APPLICATION NUMBER: US/09/788,133
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-788-133-2

Query Match 18.5%; Score 395; DB 10; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

Qy 107 LSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRKTS-ISLVIFHTNLAIADLLFCVTLPLFKI 165

Db 29 LENMLFASFYLLDFILALVGNLTALWLFIRDHKSQTPANVFLMHLAVADLSCLVLPTRL 88

Qy 166 AYHLNGNNWVFGVCMCRITTVFVGNMYCAIILITCMGINRYLATATAHPFTYQKLPKRSFS 225
Db 89 VYHFGNHNWPFGEIACRLTGLFGLYNNMYSIYFLTCISADRFALVHPVKSLLRRPLA 148

Qy 226 LLMCGIYVMVFLYMLPFVILKQYHLVHSEITTCCHDVVDACESPSSFRFYFVSLAFFG 285
Db 149 HLACAFLLWVVAVAMAPLLVSPQ-----TVQTNHTVCLQLYREKASHHALVSLA-VA 200

Qy 286 FLIPFVIIIFCYTTLIHKLK-----KORIWLGYIKAVLLILVIFTICFAPTNIILVIHHA 341
Db 201 FTFFPITVTCLLIIRSLRQGLRVEKRLKTKAVRMIAIIVLFLVCFVYVHNRSVYVL 260

Qy 342 NYHYHTDS-----LYFWYLIALCLGSLNSCLDPLFYVMSK 378
Db 261 HYRSHGASCATORILALANRITSCLTSLNGALDPIMYFFVAE 302

RESULT 19

US-10-092-135-4
; Sequence 4, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 362
; TYPE: PRT
; ORGANISM: MELEAGRIS GALLOPAVO
US-10-092-135-4

Query Match 18.5%; Score 395; DB 9; Length 362;
Best Local Similarity 28.0%; Pred. No. 1.5e-26;
Matches 90; Conservative 64; Mismatches 122; Indels 46; Gaps 10;

Qy 73 GMTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLW 132
Db 21 GWAAGNASTKC-----SLTKTGQFYFLPTVYILVFTITGLGNSVAIW 63

Qy 133 KLSLRTK-----SISLVIFHTNLAIADLLFCVTLPLFKIAYHLNGNNWVFGVCMCRITTVFY 189
Db 64 MEVFNHRPWSGISVYMF--NLALADFLYVTLTLPALIFYFNKTDWTFDGMCKLQRFIFH 121

Qy 190 GNMCAIILITCMGINRYLATATAHPFTYQKLPKRSPLMCGIYVMVFLYMLPFVILKQE 249
Db 122 VNLVGSILFELTCLISVHRYTGVVHPLKSLGRLKKNAVYVSSLVWALVAVIAP-ILFYSG 180

Qy 250 YHLVHSEITTCCHDVVDACESPSSFRFYFVSLAFFGFLIPFVIIIFCYTTLIHKLKSKD- 308
Db 181 TGVRRNKITCYDIT-ADVLRSY-FVYSMCTTVFMFCIPFVILGCGYGLIVKALIVKDL 238

Qy 309 -----RIWLGYIKAVLLILVIFTICFAPTNIILVIH-HANYYYHN-----TSLYF 353
Db 239 DNSPLRKRKIYL-----VIIVLTFAVSILPFHVMKTLNRLARLDFQTPMCAFNKDYA 293

Qy 354 MYLIALCLGSLNSCLDPLFYFV 375
Db 294 TYQVTRGLASLNSCVDPILYFL 315

RESULT 20

```
US-10-311-956-4
; Sequence 4, Application US/10311956
; Publication No. US20030109482A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; APPLICANT: Bull, Christof
; TITLE OF INVENTION: REGULATION OF HUMAN P2Y8-LIKE G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: Lio095 foreign countries
; CURRENT APPLICATION NUMBER: US/10/311,956
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/214,257
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 60/254,878
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-311-956-4

Query Match      18.4%; Score 393; DB 9; Length 537;
Best Local Similarity 28.5%; Pred. No. 3.5e-26;
Matches 95; Conservative 62; Mismatches 116; Indels 60; Gaps 11;

Qy 87 EDSISTLHVNNATIGYRSSL-----STQ-----VIPAIYIILLFVVGVPNSI 128
Db 3 EDIATSYPTFLTPPYLPKMLNLTNDTEICVDFDEGFKFLLLPVSYSAVFWGLPLNI 62

Qy 129 VTWLKLSRYSIS-LVIFHTNLAIDLFCVTLFPFKIAHYHLNNGNNVFGVWMCRTTVV 187
Db 63 AAMWIFIAKRPWNPPTVYMFNLALSDTLVLSLPTLVYVYADKNNWPFGEVLCKLVRFL 122

Qy 188 FYGNMCAIILITCMGINRYLATAPETYOKLPKRSFLMCGIWMVWVLYMLPFVILX 247
Db 123 FYANLYSSILFTICISVHRVYGVCHPITSURRMAKHAIVICALVLSVLTCLVFNLI-- 180

Qy 248 QEYHLVHSEI--TTCHDVVDACESSPSFRFY--YFVSLAFFGLFPFVILFVITLILHK 303
Db 181 --FVTVSPKVNKTCHDTT-----RPEDFARVVESTAIMCLLFGICLLIAGCYGLMTR 234

Qy 304 LKS-----KDRWLGVYKAVLILVIFTICFAPTNIILVIHANYYY-----345
Db 235 LMKPIVSGNQOTLPSYKKR-----SIKTIIFWIAFAICFMPFHITRTL-----YYARLLG 286

Qy 346 ---HNTDSLFWMYLIALCLGSLNSCLDPLFYFV 375
Db 287 IKCYALNVINVTYKVTREPLANSCLDPIFYFL 319

RESULT 21
US-10-251-385-182
; Sequence 182, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; FILE REFERENCE: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-10-251-385-182
Query Match      18.2%; Score 389; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 4.5e-26;
Matches 88; Conservative 52; Mismatches 124; Indels 18; Gaps 5;

Qy 107 LSTQVTPAIYIILLFVVGVPNSIVTLWLKLSRYSIS-LVIFHTNLAIDLFCVTLFPFKI 165
Db 29 LENMLFASFYLLDFILALVGNLTALWLFTRDHKSGTTPANVFLMHLAVADLSCVLVLPTRL 88

Qy 166 AYHLNGNNVFGVWMCRTTVVFGYNNMCAIILITCMGINRYLATAPETYOKLPKRSFS 225
Db 89 VTHFSGNHPFGEIACRLTGFLFYLNMYASIIFLTICISADRFIAIVHPVKSLLKRLPLVA 148

Qy 226 LLMCGIWMVWVLYMLPFVILKQYHLVHSEITTCCHDVVDACESSPSFRFYFVSLAFFG 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVVCLQLYREKASHALVSLA-VA 200

Qy 286 FLIPFVILFICYTTLIHKLS-----KDRWLGVYKAVLILVIFTICFAPTNIILVIHHA 341
Db 201 FTFPFITTVCYLLIIRSLRQGLRVEKRLKTRAKRMIAIVLAIFLVCVFPYHVNRSVYVL 260

Qy 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 261 HYRSHGASCATORILALANRITSCLTSLNGALDPIIMYFFVAE 302

RESULT 22
US-09-779-679-25
; Sequence 25, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Vellizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A.
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sa
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 366
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-779-679-25

Query Match      18.2%; Score 388; DB 9; Length 366;
Best Local Similarity 28.3%; Pred. No. 6e-26;
Matches 99; Conservative 60; Mismatches 135; Indels 56; Gaps 10;

Qy 42 SDNSAKPTLTIKSPNG--GPQTFEFPLSDIEGWTGATTIKAECPEDSISTLHVNA 98
Db 23 SDSS-----QSNGLVAPPGLITNPSLATAE-----QCQE-----54
Qy 99 TIGYLRSLSTQVPAIYIILFVVGVPNSIVTLWKLRLTKS-ISLVIFHTNLAIADLLF 157
Db 55 -----TPLENMLFASFYLLDFILALVGNLTALWLFIRDHKSGTPANVFLMHLAVADLSC 108
Qy 158 CVTLPPFKIAHLNGNNVFGVEVMCRITTVFVGNMYCAILLICMGINRYLATAHPTYQ 217
Db 109 VLVLPTLLVHFSGNHPFGEIACRLTGLFLYLNMYASIVFLTCISADRELA-VHPVKSL 167
Qy 218 KLPKRSFSLLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCCHDVVDACESPSPRFY 277
Db 168 KLRPLVAHLACALWVAVAMAPLLVSPQ-----TVQTNHTVVCLOLYREKASHHA 220
Qy 278 FVSLAFGFLIPFVIIIFCYTTLTHKLK-----KDRWLGYKAVLILVIFTCFAPT 333
Db 221 LVSLA-VAFTPTFTTTCYLLIIRSLRQGLRVEKRLKTRAVRMIAIVLAIFLVCFVPYH 279
Qy 334 IILVIHANYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 280 VNRSVYVLYRSHGASCATORILANRITSCLTSLNGALDPIIMYFFVAE 329

RESULT 23
US-09-826-791-2
; Sequence 2, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1e1 Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US/09/826,791
; PRIOR FILING DATE: 2001-04-05
; PRIOR FILING DATE: 2000-04-05
; PRIOR FILING DATE: 2000-04-05
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 2
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-791-2

Query Match      17.7%; Score 378.5; DB 10; Length 330;
Best Local Similarity 28.9%; Pred. No. 3.6e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

Qy 86 PEDSISTLHVNNATIGYLRSSLSTQVPAIYIILFVVGVPNSIVTLWK-LSLRTKSISLV 144
Db 3 PNGTFSNNNSRNTI-----ENFKREFPIVYLIIFFWGLNGLSIYVFLQPKKSTSVN 58
Qy 145 IFHTNLAIADLLFCVTLPPFKIAYHLNGNNVFGVEVMCRITTVFVGNMYCAILLITCMGI 204
Db 59 VFMLNLAIADLLFISTLPFRADYLRGNSWIFGLACRIMSYSLYVNNYSIYFLTVLSV 118
Qy 205 NRYLATAHPFTYQKLPKRSFSLLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCCHDV 264
Db 119 VRFLAMVHPFRLHVTISRSAWILCGIWIIL----MASSIMLDGSGEQSGSVTSCLEL- 174
Qy 265 DACESPSPRFYFVSL-----AFPGFLIPFVIIIFCVTLTHLKLK-----KDRWL 313
Db 175 -----NLYKIAKLTQMNIALVVGCLLPFTLSICVLLIIRVLLKVEVPESGLRVS 225

; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Vellizar I
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
```

```
; ORGANISM: Homo sapiens
US-09-828-478-2

Query Match      17.7%; Score 378.5; DB 9; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

Qy 86 PEDSISTLHVNNATIGYLRSSLSTQVPAIYIILFVVGVPNSIVTLWK-LSLRTKSISLV 144
Db 19 PNGTFSNNNSRNTI-----ENFKREFPIVYLIIFFWGLNGLSIYVFLQPKKSTSVN 74
Qy 145 IFHTNLAIADLLFCVTLPPFKIAYHLNGNNVFGVEVMCRITTVFVGNMYCAILLITCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYLRGNSWIFGLACRIMSYSLYVNNYSIYFLTVLSV 134
Qy 205 NRYLATAHPFTYQKLPKRSFSLLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCCHDV 264
Db 135 VRFLAMVHPFRLHVTISRSAWILCGIWIIL----MASSIMLDGSGEQSGSVTSCLEL- 190
Qy 265 DACESPSPRFYFVSL-----AFPGFLIPFVIIIFCVTLTHLKLK-----KDRWL 313
Db 191 -----NLYKIAKLTQMNIALVVGCLLPFTLSICVLLIIRVLLKVEVPESGLRVS 241
Qy 314 YIKA-----VLLILVIFTCFAPTNIILVIHANYYYH-NTDSLIFMYLIALCLGSLNSCLD 369
Db 242 HRKALTITITLITLIFELCELPVHTLRTVHLTTWKVGLCKDRHLKALVITLALAAANACFN 301
Qy 370 PFLYF 374
Db 302 PLYY 306

RESULT 25
US-09-779-679-2
; Sequence 2, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Vellizar I
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
```



```
US-09-826-791-6
; Sequence 6, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-791-6

Query Match      17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

QY 86 PEDSISTLHVNNTATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWK-LSLRTKTSISLV 144
Db 19 PNGTFSNNNSRNCIT-----ENFKREFFPIVLIIFWVGVLGNGLSIYVFLQPYKKTSV 74

QY 145 IFHTNLAIADLLFCVTLFPFKIAIYHLNGNNWVGEVCMCRITTVVFGYGNMYCAIILTCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYILRGSNWIFGDLACRIMSYSLYVNMYSIYFLTVLSV 134

QY 205 NRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCDDVV 264
Db 135 VRFLAMVHPFRLHVTIRSAMILCGIWIIL---MASSIMLLDSGSEQNGSVTSCLEL- 190

QY 265 DACESPSSFRFYFVSL-----AFFGFLIPFVIIIFCYTTLIHKLK-----KDRIWLG 313
Db 191 -----NLYKIAKIQTNMYIALVVGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVS 241

QY 314 YIKA---VLLILVIFTICFAPTNIIIVIHANYHH-NTDSLYFMYLIAICLGLSCLD 369
Db 242 HRKALTTIIITLIIFLCLFPLPYHTLRTVHLTWKVGCKDRHLKALVITLALAAANACFN 301

QY 370 PFLYF 374
Db 302 PLYY 306

RESULT 28
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 7
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-7

Query Match      17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

QY 86 PEDSISTLHVNNTATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWK-LSLRTKTSISLV 144
Db 19 PNGTFSNNNSRNCIT-----ENFKREFFPIVLIIFWVGVLGNGLSIYVFLQPYKKTSV 74

QY 145 IFHTNLAIADLLFCVTLFPFKIAIYHLNGNNWVGEVCMCRITTVVFGYGNMYCAIILTCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYILRGSNWIFGDLACRIMSYSLYVNMYSIYFLTVLSV 134

QY 205 NRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCDDVV 264
Db 135 VRFLAMVHPFRLHVTIRSAMILCGIWIIL---MASSIMLLDSGSEQNGSVTSCLEL- 190

QY 265 DACESPSSFRFYFVSL-----AFFGFLIPFVIIIFCYTTLIHKLK-----KDRIWLG 313
Db 191 -----NLYKIAKIQTNMYIALVVGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVS 241

QY 314 YIKA---VLLILVIFTICFAPTNIIIVIHANYHH-NTDSLYFMYLIAICLGLSCLD 369
Db 242 HRKALTTIIITLIIFLCLFPLPYHTLRTVHLTWKVGCKDRHLKALVITLALAAANACFN 301

QY 370 PFLYF 374
Db 302 PLYY 306

RESULT 28
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 7
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-7

Query Match      17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;
```

```
QY 86 PEDSISTLHVNNTATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWK-LSLRTKTSISLV 144
Db 19 PNGTFSNNNSRNCIT-----ENFKREFFPIVLIIFWVGVLGNGLSIYVFLQPYKKTSV 74

QY 145 IFHTNLAIADLLFCVTLFPFKIAIYHLNGNNWVGEVCMCRITTVVFGYGNMYCAIILTCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYILRGSNWIFGDLACRIMSYSLYVNMYSIYFLTVLSV 134

QY 205 NRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCDDVV 264
Db 135 VRFLAMVHPFRLHVTIRSAMILCGIWIIL---MASSIMLLDSGSEQNGSVTSCLEL- 190

QY 265 DACESPSSFRFYFVSL-----AFFGFLIPFVIIIFCYTTLIHKLK-----KDRIWLG 313
Db 191 -----NLYKIAKIQTNMYIALVVGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVS 241

QY 314 YIKA---VLLILVIFTICFAPTNIIIVIHANYHH-NTDSLYFMYLIAICLGLSCLD 369
Db 242 HRKALTTIIITLIIFLCLFPLPYHTLRTVHLTWKVGCKDRHLKALVITLALAAANACFN 301

QY 370 PFLYF 374
Db 302 PLYY 306

RESULT 29
US-09-866-230-9
; Sequence 9, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Ma
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 9
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-9

Query Match      17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

QY 86 PEDSISTLHVNNTATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWK-LSLRTKTSISLV 144
Db 19 PNGTFSNNNSRNCIT-----ENFKREFFPIVLIIFWVGVLGNGLSIYVFLQPYKKTSV 74

QY 145 IFHTNLAIADLLFCVTLFPFKIAIYHLNGNNWVGEVCMCRITTVVFGYGNMYCAIILTCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYILRGSNWIFGDLACRIMSYSLYVNMYSIYFLTVLSV 134

QY 205 NRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEHVHSEITTCDDVV 264
Db 135 VRFLAMVHPFRLHVTIRSAMILCGIWIIL---MASSIMLLDSGSEQNGSVTSCLEL- 190

QY 265 DACESPSSFRFYFVSL-----AFFGFLIPFVIIIFCYTTLIHKLK-----KDRIWLG 313
Db 191 -----NLYKIAKIQTNMYIALVVGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVS 241

QY 314 YIKA---VLLILVIFTICFAPTNIIIVIHANYHH-NTDSLYFMYLIAICLGLSCLD 369
Db 242 HRKALTTIIITLIIFLCLFPLPYHTLRTVHLTWKVGCKDRHLKALVITLALAAANACFN 301

QY 370 PFLYF 374
Db 302 PLYY 306
```

```
RESULT 30
US-09-828-478-5
; Sequence 5, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828.478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-5

Query Match      17.7%; Score 378; DB 9; Length 337;
Best Local Similarity 32.6%; Pred. No. 4e-25;
Matches 98; Conservative 55; Mismatches 128; Indels 20; Gaps 9;

QY 88 DSISTLHVNNTATGCLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--KSISLVI 145
Db 2 DETGNLTSSATCHDTIDDFRNQVYSTLYSMISVGVFGNGFVLYVL-IKTYHKSAFQV 60
QY 146 FHTNLAIADLLFCVTLPLFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILILTCMGIN 205
Db 61 YMINLAVADLLCVCTLPLRWVYVYHKGIMLFGDFCLRLSTYALYVNLVYCSIFFTWMSFF 120
QY 206 RYLATAHPFTYQKLPKRFSLLMCGIWMVFLYMLPFVILKQEHVHLVHSEITTCHEVDVD 265
Db 121 RCIAIVFPQVNLVTKARFVCGIWFIVLTSSPFLMAKPKQD--EKNTKCEPPQ 178
QY 266 ACESPSFRFYFVSLAFAFGFLIPFVILFCYTTLIHLK--KSKDRWLWGIYKAVLLIV 323
Db 179 DNQTKNHLVHLHYVSL-FVGFIIPFVILVVCYTMIIITLLKSKMKNLSSHKAIGMIMV 237
QY 324 I---FTICFAPTNIILVIHANYVHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
Db 238 VTAFLVSEMPHYHIQRTIH--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
QY 374 F 374
Db 296 F 296

Query Match      17.7%; Score 378; DB 9; Length 337;
Best Local Similarity 32.6%; Pred. No. 4e-25;
Matches 98; Conservative 55; Mismatches 128; Indels 20; Gaps 9;

QY 88 DSISTLHVNNTATGCLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--KSISLVI 145
Db 2 DETGNLTSSATCHDTIDDFRNQVYSTLYSMISVGVFGNGFVLYVL-IKTYHKSAFQV 60
QY 146 FHTNLAIADLLFCVTLPLFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILILTCMGIN 205
Db 61 YMINLAVADLLCVCTLPLRWVYVYHKGIMLFGDFCLRLSTYALYVNLVYCSIFFTWMSFF 120
QY 206 RYLATAHPFTYQKLPKRFSLLMCGIWMVFLYMLPFVILKQEHVHLVHSEITTCHEVDVD 265
Db 121 RCIAIVFPQVNLVTKARFVCGIWFIVLTSSPFLMAKPKQD--EKNTKCEPPQ 178
QY 266 ACESPSFRFYFVSLAFAFGFLIPFVILFCYTTLIHLK--KSKDRWLWGIYKAVLLIV 323
Db 179 DNQTKNHLVHLHYVSL-FVGFIIPFVILVVCYTMIIITLLKSKMKNLSSHKAIGMIMV 237
QY 324 I---FTICFAPTNIILVIHANYVHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
Db 238 VTAFLVSEMPHYHIQRTIH--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
QY 374 F 374
Db 296 F 296

RESULT 31
US-09-866-230-8
; Sequence 8, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866.230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-8
```

```
Query Match      17.7%; Score 378; DB 10; Length 337;
Best Local Similarity 32.6%; Pred. No. 4e-25;
Matches 98; Conservative 55; Mismatches 128; Indels 20; Gaps 9;

QY 88 DSISTLHVNNTATGCLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--KSISLVI 145
Db 2 DETGNLTSSATCHDTIDDFRNQVYSTLYSMISVGVFGNGFVLYVL-IKTYHKSAFQV 60
QY 146 FHTNLAIADLLFCVTLPLFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILILTCMGIN 205
Db 61 YMINLAVADLLCVCTLPLRWVYVYHKGIMLFGDFCLRLSTYALYVNLVYCSIFFTWMSFF 120
QY 206 RYLATAHPFTYQKLPKRFSLLMCGIWMVFLYMLPFVILKQEHVHLVHSEITTCHEVDVD 265
Db 121 RCIAIVFPQVNLVTKARFVCGIWFIVLTSSPFLMAKPKQD--EKNTKCEPPQ 178
QY 266 ACESPSFRFYFVSLAFAFGFLIPFVILFCYTTLIHLK--KSKDRWLWGIYKAVLLIV 323
Db 179 DNQTKNHLVHLHYVSL-FVGFIIPFVILVVCYTMIIITLLKSKMKNLSSHKAIGMIMV 237
QY 324 I---FTICFAPTNIILVIHANYVHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
Db 238 VTAFLVSEMPHYHIQRTIH--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
QY 374 F 374
Db 296 F 296

RESULT 32
US-10-092-135-7
; Sequence 7, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092.135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 373
; TYPE: PRT
; ORGANISM: RATTUS NORVEGICUS
US-10-092-135-7

Query Match      17.7%; Score 377.5; DB 9; Length 373;
Best Local Similarity 29.3%; Pred. No. 5e-25;
Matches 86; Conservative 65; Mismatches 110; Indels 33; Gaps 9;

QY 103 LRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLTK---SISLVIFHTNLAIADLLFCV 159
Db 45 IKTGFGFYLLPAVYIILVFIIFGLGNSVAIWMFVFMKPSGSGSVYWF--NLADLFLYVL 102
QY 160 TLFPKIAIYHLNGNNWVFGVMCRITTVFYGNMYCAILILTCMGINRYLATATAHPFTYQKL 219
Db 103 TLPALIFYFNKTDWIFGDVCMCKLQRFIFHVNLYGSILFLTCISAHRYSGVYVPLKSLGR 162
QY 220 PKRFSLLMCGIWMVFLYMLPFVILKQEHVHLVHSEITTCHEVDVDACSPSFRFYFV 279
Db 163 LKKNAIYVSVLWLIWVAISP-ILFYSGTGIRKNTVTCYD---STSDYLRSYFIY 217
QY 280 SL--AFEGFLIPFVILFCYTTLIHLKSKD-----RIWGIYKAVLLILVIFTIC 328
Db 218 SMCTTVAMFCIPLVLILGCGIVRALIYKDLNDSPLRRKSIYL-----VIIIVTFAVS 272
```


Db 179 DNQTKHVLVHVSU-FVGGIIPFVIIIVCYTMIITLLKSKMKNLSHKKAIGMIV 237
Qy 324 I---FTICFAPTNIILVIHANYHYN---TDSLXFM---YLIACLGSLNSCLDPFLY 373
Db 238 VTAFLVSPFYHQIHTI--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
Qy 374 F 374
Db 296 F 296
RESULT 35
US-10-092-135-6
; Sequence 6, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
; FILE REFERENCE: D0134.NP
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 373
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-092-135-6
Query Match 17.6%; Score 376.5; DB 9; Length 373;
Best Local Similarity 27.0%; Pred. No. 6.2e-25;
Matches 95; Conservative 72; Mismatches 130; Indels 55; Gaps 12;
Qy 46 AKPTLTITKSFNGGQNTFEFPPLSDIEGTGATTIKAECPEDSISTLHVNNATIGYLRL 104
Db 8 AVPNGTDAFLAGPSS-----WGNSTVA-----STAAVSSSFKCALTK 46
Qy 105 SSLSQTVPIAIYILLFVGVPSNIVTLWLSLRTK---SISLVIFHTNLAIADLLFCVTL 161
Db 47 TGFOFYVLPVAYILVFIIGLNSVAIWMFVFMKPSGIVYMF--NLALADFLYVLT 104
Qy 162 PFKIAYHLNGNWNWGEVGMCRITTVFYGNMYCAIILITCMGINRYLATAHPTTYQKLPK 221
Db 105 PALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFTICSAHRYSGVYVYPLKSLGRK 164
Qy 222 RSFLLMCGIWMVFLYMLPFVILKQYHLVHSEITTDHVDVADACESPSFRFYFVSL 281
Db 165 KKAICISVLWLVVVAISP-ILFYSGTGVRKNKTTICYDTT-----SDEYLSRYFIYSM 219
Qy 282 --AFFGFLIPFVIIIFCYTTLIHKLKSKD-----RIWLGVIKAVLLIVIFTCFAPTNIIL 330
Db 220 CTTVAMFCVPLVILGCVGLVIRALIVKLDNSPLRKSIVL-----VIIVLVFAVSYI 274
Qy 331 PTNIILVIH-HANYHYN-----TDSLXFMYLIALCLGSLNSCLDPFLYFV 375
Db 275 PFVWMTNLRARLDFTQPAWCAFNDRVYATYQVTRGLASLNSCVDPLVFL 326
RESULT 36
US-10-092-135-5
; Sequence 5, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/09/943,718
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>

; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 373
; TYPE: PRT
; ORGANISM: BOS TAURUS
US-10-092-135-5
Query Match 17.4%; Score 372; DB 9; Length 373;
Best Local Similarity 27.8%; Pred. No. 1.5e-24;
Matches 94; Conservative 63; Mismatches 139; Indels 42; Gaps 10;
Qy 59 PQNTFEFPLSDIEGTGATTIKAECPEDSISTLHVNNATIGYLSSLSQTVPIAIYL 118
Db 10 PNGTDTAFLADPGSPMGNSTVTSTAASVAFKCALTKTGFYYL-----PAVYIL 60
Qy 119 LFVGVPSNIVTLWLSLRTK---SISLVIFHTNLAIADLLFCVTLPFKIAYHLNGNWN 175
Db 61 VFIIGLNSVAIWMFVFMKPSGIVYMF--NLALADFLYVLTLPALIFYFNKTDWI 118
Qy 176 FGEVMCRITTVFYGNMYCAIILITCMGINRYLATAHPTTYQKLPKRSFSLMCGIWMV 235
Db 119 FGDAMCKLQRFIFHVNLYGSILFTICSAHRYSGVYVYPLKSLGRKKNNAVYISVLWLI 178
Qy 236 VFYMLPFVILKQYHLVHSEITTDHVDVADACESPSFRFYFVSL--AFFGFLIPFVII 293
Db 179 VVVGISP-ILFYSGTGVRKNKTTICYDTT-----SDEYLSRYFIYSCTTVAMFCVPLVLI 233
Qy 294 IFCYTTLIHKLKSKD-----RIWLGVIKAVLLIVIFTCFAPTNIILVIH-HANY 343
Db 234 LGCYGLVIRALIVKLDNSPLRKSIVL-----VIIVLVFAVSYIPFVWMTNLRARL 288
Qy 344 YHYN-----TDSLXFMYLIALCLGSLNSCLDPFLYFV 375
Db 289 DFQTPMCAFNDRVYATYQVTRGLASLNSCVDPLVFL 326
RESULT 37
US-09-943-718-4
; Sequence 4, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Hufine, Constance F.
; Rossi, Devora L.
; Capone, Myriam
; Hedrick, Joseph A.
; Vicari, Alain
; Gorman, Daniel M.
; Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/943,718
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>


```
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
; US-10-270-144-2

Query Match      17.3%; Score 369.5; DB 9; Length 337;
Best Local Similarity 29.6%; Pred. No. 2.2e-24;
Matches 87; Conservative 58; Mismatches 108; Indels 41; Gaps 8;

Qy 107 LSTQVIPAIYILLFVGVGPNIVTL-----WKLSLRKTSISLVIFHTNLAIADLLF 157
Db 31 LKMYLPVIYGIIFLVGPGNAVVIYSTYIFKMRPWK-----SSTIIMLNACTDLY 82

Qy 158 CVTLPPFKIAYHLNGNNVFGVMCRITTVVFYGNMYCAILITCMGINRYLATAHPPTYQ 217
Db 83 LTSLPFLIHYYASGENWIFGDFMCKFIRFSHFNLYSILFLCFSIFRYCVIHPMSCF 142

Qy 218 KLPKRSFLLMCGIWMVWFLYMLPFVILKQEVHLVHSEITTCDDVVVDACESPSSPFYY 277
Db 143 SIHKTRCAVVACAVWIIISLVAVIPMTFLTSTN--RTNRSACLDLTSSDEL-NTIKWYN 199

Qy 278 FVSLAFPGELIPFVIIIFCYTTLHK-----LKSXDRIWGLGVYIKAVLLILVIFTI 327
Db 200 LI-LTATTFCPLVIVTLCTYIIHTLTHGLQDSCLEKQKAR-----RTILLLLAFYV 252

Qy 328 CFAPTNILVHHANYHHNT-----DSLRYFMYLIACLSINSLCDPFLYFVMS 377
Db 253 CFLPFLILRVIRIESRLISCSIEHQIHEAYIVSRPLAALNTFGNLLLYVWVS 306
```

Search completed: June.24, 2003, 12:18:35
Job time : 35.4601 secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:53:16 ; Search time 14.4822 Seconds
(without alignments)
1165.628 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136

Sequence: 1 TLYTQHPVAGSQDINKML.....AMARPLXPRRDIEDIHAW 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1940	90.8	369	1 PAR3_MOUSE	O08675 mus musculu
2	1645.5	77.0	368	1 PAR3_RAT	Q920e1 rattus norv
3	1425.5	66.7	374	1 PAR3_HUMAN	O00254 homo sapien
4	582.5	27.3	420	1 PAR1_XENLA	P47749 xenopus lae
5	581.5	27.2	399	1 PAR2_MOUSE	P55086 mus musculu
6	569	26.6	397	1 PAR2_HUMAN	P55085 homo sapien
7	564.5	26.4	397	1 PAR2_RAT	Q63845 rattus norv
8	540.5	25.3	430	1 PAR1_MOUSE	P30558 mus musculu
9	522.5	24.5	428	1 PAR1_CRILLO	O00991 cricetus
10	512.5	24.0	432	1 PAR1_RAT	P26824 rattus norv
11	508	23.8	425	1 PAR1_PAPHA	P56488 papio hamad
12	507.5	23.8	395	1 PAR4_HUMAN	Q9610 homo sapien
13	501	23.5	425	1 PAR1_HUMAN	P25116 homo sapien
14	486	23.2	395	1 PAR4_RAT	Q920e0 rattus norv
15	494	23.1	396	1 PAR4_MOUSE	O88634 mus musculu
16	431.5	20.2	361	1 EBI2_HUMAN	P32249 homo sapien
17	399	18.7	362	1 P2YR_CHICK	P14996 gallus gall
18	398.5	18.7	367	1 GP17_HUMAN	Q13304 homo sapien
19	395	18.5	345	1 CLT2_PIG	Q95803 sus scrofa
20	395	18.5	362	1 P2YR_MEUGA	P49652 meleagris g
21	393	18.4	537	1 P2YR_XENLA	P79928 xenopus lae
22	387.5	18.1	370	1 P2Y9_HUMAN	Q99677 homo sapien
23	383	17.9	308	1 P2Y5_CHICK	P32250 gallus gall
24	382.5	17.9	309	1 CLT2_RAT	Q924t9 rattus norv
25	378.5	17.7	346	1 CLT2_HUMAN	Q9875 homo sapien
26	378.5	17.7	373	1 P2Y2_MOUSE	P35383 mus musculu
27	378	17.7	337	1 CLT1_HUMAN	Q9Y271 homo sapien
28	377.5	17.7	361	1 P2Y4_RAT	O35811 rattus norv
29	377.5	17.7	373	1 P2YR_RAT	P49651 rattus norv
30	376.5	17.6	373	1 P2YR_HUMAN	P47900 homo sapien
31	372.5	17.4	361	1 P2Y4_MOUSE	Q9JJ57 mus musculu
32	372	17.4	373	1 P2YR_BOVIN	P48042 bos taurus
33	367	17.2	381	1 GP34_HUMAN	Q9upc5 homo sapien

RESULT 1

ID	PAR3_MOUSE	STANDARD;	PRT;	369 AA.
AC	O08675;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-like 2) (Coagulation factor II receptor-like 2).			
DE	like 2) (Coagulation factor II receptor-like 2).			
GN	P2RL2 OR PAR3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RX	MEDLINE=97242411; PubMed=9087410;			
RA	Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W.,			
RA	Timmons C., Tram T., Coughlin S.R.;			
RT	"Protease-activated receptor 3 is a second thrombin receptor in humans."			
RL	Nature 386:502-506 (1997).			
CC	FUNCTION: High affinity receptor for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis. May play a role in platelets activation.			
CC	SUBCELLULAR LOCATION: Integral membrane protein.			
CC	PTM: A proteolytic cleavage generates a new amino terminus that functions as a tethered ligand.			
CC	SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; U92972; AAC53137.1; --			
DR	HSSP; P34996; 1DDD.			
DR	MGD; MGI:1298208; P2rl2.			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	Pfam; PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCRHHODPSN.			
DR	PROSITE; PS00237; G-PROTEIN RECP F1.1; FALSE_NEG.			
DR	PROSITE; PS00262; G-PROTEIN_RECP_F1.2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Blood coagulation.			
FT	SIGNAL	1	19	POTENTIAL.
FT	PROPEP	20	37	REMOVED FOR RECEPTOR ACTIVATION (BY SIMILARITY).
FT	CHAIN	38	369	PROTEINASE ACTIVATED RECEPTOR 3.
FT	DOMAIN	38	93	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	94	119	1 (POTENTIAL).
FT	DOMAIN	120	127	CYTOPLASMIC (POTENTIAL).

ALIGNMENTS

FT TRANSMEM 128 147 2 (POTENTIAL).
 FT DOMAIN 148 166 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 167 188 3 (POTENTIAL).
 FT DOMAIN 189 205 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 206 229 4 (POTENTIAL).
 FT DOMAIN 230 258 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 260 279 5 (POTENTIAL).
 FT DOMAIN 280 296 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 297 321 6 (POTENTIAL).
 FT DOMAIN 322 335 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 336 360 7 (POTENTIAL).
 FT DOMAIN 361 369 CYTOPLASMIC (POTENTIAL).
 FT SITE 37 38 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 165 244 BY SIMILARITY.
 FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 369 AA; 41707 MW; F4166BF766D073DB CRC64;

Query Match 90.8%; Score 1940; DB 1; Length 369;
 Best Local Similarity 100.0%; Pred. No. 2.9e-101;
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 MKLILVAAGLLPLPVTVCOSGINVSDNSAKPTLTITKSPNGGPONTFFEPFLSDIEGWTG 76
 DB 1 MKLILVAAGLLPLPVTVCOSGINVSDNSAKPTLTITKSPNGGPONTFFEPFLSDIEGWTG 60

QY 77 ATTITKAECPEDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNSIVTLWKL 136
 DB 61 ATTITKAECPEDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNSIVTLWKL 120

QY 137 RTKSISLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNNVFGVCMCRITTVFVGNMYCAI 196
 DB 121 RTKSISLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNNVFGVCMCRITTVFVGNMYCAI 180

QY 197 LILTCGGINRYLATAHPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSE 256
 DB 181 LILTCGGINRYLATAHPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSE 240

QY 257 ITTCHDVVDACESPSSFRFYVSLAFFGFLIPFVILIIICVYTLIHLKSKDRIMWLYIK 316
 DB 241 ITTCHDVVDACESPSSFRFYVSLAFFGFLIPFVILIIICVYTLIHLKSKDRIMWLYIK 300

QY 317 AVLLILVIFTCFAPTNILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLFYFM 376
 DB 301 AVLLILVIFTCFAPTNILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLFYFM 360

QY 377 SKYVDQLNP 385
 DB 361 SKYVDQLNP 369

RESULT 2

PAR3_RAT ID PAR3_RAT STANDARD; PRT; 368 AA.
 AC Q920E1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-like 2) (Coagulation factor II receptor-like 2).
 GN F2RL2 OR PAR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
 RA Chien E.K., Marietti S., Mendoza J., Phillippe M.;
 RT "Cloning of the rat protease activated receptor isoforms 3 and 4.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- PTM: A proteolytic cleavage generates a new amino terminus that functions as a tethered ligand.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF310076; AAL36789.1; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Blood coagulation.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 37 REMOVED FOR RECEPTOR ACTIVATION (BY SIMILARITY).
 FT CHAIN 38 368 PROTEINASE ACTIVATED RECEPTOR 3.
 FT DOMAIN 38 93 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 94 114 1 (POTENTIAL).
 FT DOMAIN 115 123 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 124 144 2 (POTENTIAL).
 FT DOMAIN 145 166 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 167 187 3 (POTENTIAL).
 FT DOMAIN 188 208 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 209 229 4 (POTENTIAL).
 FT DOMAIN 230 257 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 258 278 5 (POTENTIAL).
 FT DOMAIN 279 300 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 301 321 6 (POTENTIAL).
 FT DOMAIN 322 338 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 339 359 7 (POTENTIAL).
 FT DOMAIN 360 368 CYTOPLASMIC (POTENTIAL).
 FT SITE 37 38 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 164 243 BY SIMILARITY.
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 368 AA; 41795 MW; CE0E94EDA3B80EF1 CRC64;

Query Match 77.0%; Score 1645.5; DB 1; Length 368;
 Best Local Similarity 83.5%; Pred. No. 6e-85;
 Matches 309; Conservative 31; Mismatches 25; Indels 5; Gaps 3;

QY 15 IMKILILVAAGLLPLPVTVCOSGI-NVSDNSAKPTLTITKSPNGGPONTFFEPFLSDIEG 73
 DB 1 MEMKVILVGVRLPLPVTVCOSGMKHSVDSNA---LTAESFNGN-EHSFEFPFLSDIEG 56

QY 74 WTGATTTIKACPEDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNSIVTLW 133
 DB 57 WTGATTTIKACPEESITTLHVNNATMGYLRSLSTQVIPAIIYLLFVVGVPANTVTLW 116

QY 134 LSLRTKSLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNNVFGVCMCRITTVFVGN 193
 DB 117 LSSRTKSLVIFHTNLAIADLLFCVTLPPKIAHYHLNGNDVFGVCMCRITTVFVGN 176

QY 194 CAILILTCGGINRYLATAHPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLV 253
 DB 177 CAILILTCGGINRYLATAHPFTYQKLPKRNFTLLMCGVWMVFLYMLPLAILKOEYHLV 236

QY 254 HSEITTCCHDVVDACESPSSFRFYVSLAFFGFLIPFVILIIICVYTLIHLKSKDRIMW 313
 DB 237 QGQITTCCHDVHDTCSPLPFPQFYVSLAFFGFLIPFVVSFVFCYTTLILHLKAQDKWL 296

QY 314 YIKAVLLILVIFTCFAPTNILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLY 373
 DB 297 YIKAVLLILVIFTCFAPTNILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLY 356

QY 374 FVMSKVVDQL 389


```
RESULT 4
PARI_XENLA
ID PARI_XENLA STANDARD; PRT; 420 AA.
AC F47749;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94195429; PubMed=8145852;
RA Gerszten R.E., Chen J., Ishii M., Ishii K., Nanavicz T.,
RA Turk C.W., Vu T.-K.H., Coughlin S.R.;
RT "Specificity of the thrombin receptor for agonist peptide is defined
RT by its extracellular surface.";
RL Nature 368:648-651(1994).
CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U09632; AA18498.1; -
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G-PROTEIN RECP F1.1; 1.
CC PROSITE; PS0262; G-PROTEIN RECP F1.2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Blood coagulation.
CC SIGNAL 1 20 POTENTIAL.
CC PROPEP 21 42 REMOVED FOR RECEPTOR ACTIVATION.
CC CHAIN 43 420 PROTEINASE ACTIVATED RECEPTOR 1.
CC DOMAIN 43 101 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 102 127 1 (POTENTIAL).
CC DOMAIN 128 136 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 137 156 2 (POTENTIAL).
CC DOMAIN 157 175 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 176 197 3 (POTENTIAL).
CC DOMAIN 198 217 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 218 238 4 (POTENTIAL).
CC DOMAIN 239 267 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 268 287 5 (POTENTIAL).
CC DOMAIN 288 310 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 311 333 6 (POTENTIAL).
CC DOMAIN 334 345 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 346 369 7 (POTENTIAL).
CC DOMAIN 370 420 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 295 300 POLY-SER.
CC SITE 42 43 CLEAVAGE (BY THROMBIN).
CC DISULFID 174 253 BY SIMILARITY.
CC CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC MUTAGEN 42 42 K-A: LACK OF ACTIVATION BY THROMBIN.
CC SEQUENCE 420 AA; 47435 MW; D5163F56AFE12372 CRC64;
Query Match 27.3%; Score 582.5; DB 1; Length 420;
Best Local Similarity 32.7%; Pred. No. 5.3e-26;
```

```
Matches 124; Conservative 82; Mismatches 138; Indels 35; Gaps 11;
QY 19 ILILVAAGLLPLVTVTCOSGINVSDNSAK-----PLTIKSFN--GGPONTTEEPFLSDIE 72
DB 7 LLLLLLLTLLGAMGSLCLAN---SDTQAKGHSNNMTIKTRIFDDSSSEFEIIPWDEL 63
QY 73 GWTGATTITKAECPEPDSISTL--HVNNTATIGYLRSSLSSTQVIPAIIYILLVVGVPSPN--- 127
DB 64 E-SGEGSDQAPVRSARKPIRRNITKEAEQYLSQWLTKFVPSLYTVVIVGLPLNLLA 122
QY 128 -IVTLWKLSTKTSISLVFHTNIAIADLLFCVTLFPKIAHYHLNNGNWNVFEVMCRITTV 186
DB 123 IIFLFKMKVRKPA---VVMNLNLAIDVFFVSVLPFKIAYHLSGNDWLFEGMCRIVTA 179
QY 187 VFYGNMYCAILLITCMGINRYLATAHPFYQKLPKRSFSLMCGIVVMVFLYMLPFVIL 246
DB 180 IFYCNMYCSVLLIASISVDRLAVVYPMHSLSWRTMSRAYMACSFIWLISASTIPLIVT 239
QY 247 KOEHLVHSEITTHDHDVVDACESPSFRFYFVSLAFEGFLIPFVILIIFFCYTTLHKLS 306
DB 240 EQTKIPRLDITTHDHDVLD-LKDLKDFVYVYFSSFCLLFFVFPFIITICIGIIRSLSS 298
QY 307 -----KDRILGYIKAVLLILVIFICFAPNIIIVHHANYYNNTSLYFWYLLA 358
DB 299 SSIENSKCKTRA---LFLAVVLCVFTICFGPTNVLEFLTH---YLOEANEFLYFAYILS 351
QY 359 LCLGSLNSCLDPFLYFVMS 377
DB 352 ACVGSVCCLDPLIYYAS 370
RESULT 5
PARI_MOUSE
ID PARI_MOUSE STANDARD; PRT; 399 AA.
AC P55086;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE like 1) (Coagulation factor II receptor-like 1).
DE F2RL1 OR PAR2 OR GPR11 OR GPCR11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95197620; PubMed=7890726;
RA Nystedt S., Larsson A.-K., Aaberg H., Sundelin J.;
RT "The mouse proteinase-activated receptor-2 cDNA and gene. Molecular
RT cloning and functional expression.";
RL J. Biol. Chem. 270:5950-5955(1995).
CC -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
CC G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z48043; CAA88097.1; -
CC HSSP; P34996; 1DDD.
CC MGI; MGI:101910; F2rl1.
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
```

DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 38 REMOVED FOR RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT CHAIN 39 399 PROTEINASE ACTIVATED RECEPTOR 2.
 FT DOMAIN 39 77 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 78 103 1 (POTENTIAL).
 FT DOMAIN 104 112 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 113 132 2 (POTENTIAL).
 FT DOMAIN 133 151 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 152 173 3 (POTENTIAL).
 FT DOMAIN 174 192 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 193 213 4 (POTENTIAL).
 FT DOMAIN 214 243 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 244 262 5 (POTENTIAL).
 FT DOMAIN 263 287 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 288 310 6 (POTENTIAL).
 FT DOMAIN 311 325 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 326 349 7 (POTENTIAL).
 FT DOMAIN 350 399 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 385 394 POLY-SER.
 FT DOMAIN 38 39 CLEAVAGE (BY TRYPSIN) (POTENTIAL).
 FT SITE 150 228 BY SIMILARITY.
 FT DISULFID 33 33 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 224 224 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 224 224 N-LINKED (GLNAC. . .) (POTENTIAL).
 SQ SEQUENCE 399 AA; 44752 MW; A93749425ED0B194 CRC64;

Query Match 27.2%; Score 581.5; DB 1; Length 399;
 Best Local Similarity 37.6%; Pred. No. 5.8e-26;
 Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSQTQVIAIVLLFVGVGPNIVTLKLSLRT-KSISLVIFHTNLATA 153
 DB 63 IDIFSASITGKUTTVFLPVPVIVIVGIPGNSGMALFLPRTKKHPAVYMANLAA 122
 QY 154 DLLFCVTLFPKIAHYHLGNVNVFGEVYMCRTITVVFYGNMYCAILILTCMGINRYLATAP 213
 DB 123 DLLSVIWFPLKISYHLGNVNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYVYVNP 182
 QY 214 FTYQKLPKPSFLMCGIWMVFLYMLFPVILKQBYHLVHSEITTCCHVDVACSPSSF 273
 DB 183 MGHPR-KKANIAGVSLAIWLFILFVTLPIYLMYKQTIYIPAKNITTCCHVDLPE-EVLVDG 240
 QY 274 RYFYVSLAFGFLIPFVILVFCYTTLI-----HKLKSKDRIMGLYKAVILLIV 323
 DB 241 MNYFLSLAIGVFLPFLPALUTASAYLVMIKTLRSSAMDEHSEKKRQA----IRLIITVLA 296
 QY 324 IFTICFAPTNILVHHANYHYHNTDSLXYFMYLIALCLGSLNSCLDPFLFYVMSK 378
 DB 297 MYFICFAPSNLLLVHVFYFLIKTQROSHVYALYVALCLTSLNSCIDPFVYVFSK 351

RESULT 6
 PAR2 HUMAN STANDARD; PRT; 397 AA.
 AC P55085; Q13317; Q13346;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-like 1) (Coagulation factor II receptor-like 1).
 GN P2RL1 OR PAR2 OR GPR11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96048032; PubMed=7556175;
 RA Nystedt S., Emlsson K., Larsson A.-K., Stroembeck B., Sundelin J.;
 RT "Molecular cloning and functional expression of the gene encoding the

human proteinase-activated receptor 2.";
 Eur. J. Biochem. 232:84-89(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=96177879; PubMed=8615752;
 RA Boehm S.K., Kong W., Broemme D., Smeekens S.P., Anderson D.C.,
 RA Connolly A.J., Kahn M.L., Nelken N.A., Coughlin S.R., Payan D.G.,
 RA Bunnett N.W.;
 RT "Molecular cloning, expression and potential functions of the human
 proteinase-activated receptor-2.";
 Biochem. J. 314:1009-1016(1996).
 [3]
 RP SEQUENCE OF 29-397 FROM N.A.
 RX MEDLINE=96379236; PubMed=8784787;
 RA Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A.J., Shi Y.P.,
 RA Wu R., Lin C.C., Coughlin S.R.;
 RT "Conserved structure and adjacent location of the thrombin receptor
 and protease-activated receptor 2 genes define a protease-activated
 receptor gene cluster.";
 Mol. Med. 2:349-357(1996).
 [4]
 RP SEQUENCE FROM N.A., AND VARIANTS PHE-21; GLN-270 AND ALA-291.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary, and Pancreas;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
 G proteins that stimulate phosphoinositide hydrolysis. May have a
 role in the regulation of vascular tone.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN TISSUES WITH ESPECIALLY
 HIGH LEVELS IN PANCREAS, LIVER, KIDNEY, SMALL INTESTINE, AND
 COLON. MODERATE EXPRESSION IS DETECTED IN MANY ORGANS, BUT NONE IN
 BRAIN OR SKELETAL MUSCLE.
 CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
 functions as a tethered ligand.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).

 EMBL; Z49993; CAA90290.1; -;
 EMBL; Z49994; CAA90290.1; JOINED.
 EMBL; U34038; AAB47871.1; -;
 EMBL; U36753; AAA90957.1; -;
 EMBL; AF400075; AAK77914.1; -;
 EMBL; BC012453; AAH12453.1; -;
 EMBL; BC018130; AAH18130.1; -;
 HSSP; P34996; 1DDD.
 DR Genew; HGNC:3538; P2RL1.
 DR MIM; 600933; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHOOPS
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Polymorphism.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 36 REMOVED FOR RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT CHAIN 37 397 PROTEINASE ACTIVATED RECEPTOR 2.
 FT DOMAIN 37 75 EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 76 101 1 (POTENTIAL).
FT DOMAIN 102 110 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 111 130 2 (POTENTIAL).
FT DOMAIN 131 149 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 150 171 3 (POTENTIAL).
FT DOMAIN 172 190 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 191 211 4 (POTENTIAL).
FT DOMAIN 212 241 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 242 260 5 (POTENTIAL).
FT DOMAIN 261 285 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 286 308 6 (POTENTIAL).
FT DOMAIN 309 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 347 7 (POTENTIAL).
FT DOMAIN 348 397 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 398 390 POLY-SER.
FT SITE 36 37 CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT DISULFID 148 226 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 21 21 S -> F.
FT VARIANT 270 270 /FTId=VAR_012846.
FT VARIANT 291 291 /FTId=VAR_012847.
FT VARIANT 291 291 /FTId=VAR_012848.
FT CONFLICT 138 138 G -> A (IN REF. 2).
FT CONFLICT 291 291 T -> S (IN REF. 5; AAH18130).
SQ SEQUENCE 397 AA; 44126 MW; 1A4E1D5A9B9B362B CRC64;

Query Match 26.6%; Score 569; DB 1; Length 397;
Best Local Similarity 32.7%; Pred. No. 2.8e-25;
Matches 120; Conservative 73; Mismatches 136; Indels 38; Gaps 8;

QY 22 LVAAGLLFVTVQSGINVDNSAKPTLTIKSFNGCPQNTPEEPPLSDIEGWTGATTI 81
DB 9 LLGAAILLAASCSCTIOGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KACCPEDSISTHVNNAATCYLRSSLSQVIAIYILLFVGVPSNIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDFESASVLTGKLTTFPLPIVTVIVFVGLPSNGMALWFLPRTKK 107
QY 141 ISLVIFHTNLAIADLFCTVLPKPIAHYHNGNNWGEVWCRITTVFVGNMYCAILLIT 200
DB 108 HPVAIYMANLALADLLSVFWFLPKIAHYHGNWYIYGEALCNVLIGFFYGNMYCSILEFT 167
QY 201 CMGINRYLATAHPTTYQKLPRKSFLLMCGIWMVFLYMLPVLILKQVHLVHSEITTC 260
DB 168 CLSVQRYWIVNPMGHSHR-KKANIAIGISLAIWLLLVITPIYVVKQTFIPALNITTC 226
QY 261 HDVVDACESPSSF-----RFYFVSFAFFGFLIPFVIIIFCYTTLIHKLKS-----KDRI 310
DB 227 HDVL-----PEQLLVGDMFNFLSLAIGVFLPAPLTAAYVLMIRLSSAMDENSEKK 281
QY 311 WLGYIKAVLILVIFTCTAPTNIILVIHANYHYHTDLSYPMYLIACLSGLNSCLDP 370
DB 282 RKRAIKLIVTLVLAAMYLICTPSNLLLVHVFYLIKSGQSHVYALYVALCLSLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 7
PAR2_RAT STANDARD; PRT; 397 AA.
AC Q63645;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE like 1) (Coagulation factor II receptor-like 1).
GN F2RL1 OR PAR2.
OS Rattus norvegicus (Rat).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
(1)
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Intestine, and Kidney;
MEDLINE=96358009; PubMed=8762073;
Saifeddine M., Al-Ani B., Cheng C.H., Wang L., Hollenberg M.D.;
"Rat proteinase-activated receptor-2 (PAR-2): cDNA sequence and
activity of receptor-derived peptides in gastric and vascular
tissue.";
Br. J. Pharmacol. 118:521-530(1996).
-!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
G proteins that stimulate phosphoinositide hydrolysis.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- PTM: A proteolytic cleavage generates a new amino terminus that
functions as a tethered ligand.
-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; U61373; AAC52703.1; -.
HSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 25
FT PROPEP 26 36
FT CHAIN 37 397
FT DOMAIN 37 75 PROTEINASE ACTIVATED RECEPTOR 2.
FT TRANSMEM 76 101 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 102 110 1 (POTENTIAL).
FT TRANSMEM 111 130 2 (POTENTIAL).
FT DOMAIN 131 149 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 150 171 3 (POTENTIAL).
FT DOMAIN 172 190 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 191 211 4 (POTENTIAL).
FT DOMAIN 212 241 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 242 260 5 (POTENTIAL).
FT DOMAIN 261 285 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 286 308 6 (POTENTIAL).
FT DOMAIN 309 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 347 7 (POTENTIAL).
FT DOMAIN 348 397 CYTOPLASMIC (POTENTIAL).
FT SITE 36 37 CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT DISULFID 148 226 BY SIMILARITY.
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 397 AA; 44440 MW; 27DE0C3ACABC798D CRC64;

Query Match 26.4%; Score 564.5; DB 1; Length 397;
Best Local Similarity 36.8%; Pred. No. 5e-25;
Matches 107; Conservative 69; Mismatches 106; Indels 9; Gaps 4;

QY 95 VNNATIGYRLSSLSQVIPAIVLLFVGVPSNIVTLWKLRLT-KSLSLVIFHTNLAIA 153
DB 61 VDFESAVLTGKLTTFVFLPIVIVFVGLPSNGMALWFFRTKKHPAVIYMANLALA 120
QY 154 DLLFCVTLPFKIAYHNGNNWVGEVWCRITTVFVGNMYCAILLITCGINRYLATAHP 213
DB 121 DLLSVINWFLPKISYHLHGNDWMTYDGLCKVLIGFFYGNMYCSILFCLSVQRYWVIVNP 180

```

QY 214 FTYOKLPKRSFSLMCGIWMVFLVILKQYHVLHSEITTHCHVDVACESPSSF 273
 Db 181 MHSR-KRANIAVGLSLAIWLFITVLYVMRQTYIPALNITTHCHVLP-EVLVDG 238
 QY 274 RFYVSVLAFGFLIPFVILHSEITTHCHVDVACESPSSF 327
 Db 239 MESYFLSLAIGVFLPALITASAYLMIKTLSSANDESEKKRRRAIRLITVLSWYFI 298
 QY 328 CFAPNTNIIIVHANYHTDLSYPMYIALCLGSLNSCLDPFLYFVMSK 378
 Db 299 CFAPSNVLLVWHYFLIKSQSHVYALVLCISLNSCIDPFVYFVYSK 349

RESULT 8
 ID PAR1_MOUSE STANDARD; PRT; 430 AA.
 AC P30558; P97507;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
 GN F2R OR PAR1 OR CF2R.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Coughlin S.R.;
 RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=96379236; PubMed=8784787;
 RA Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A., Shi Y.P.,
 RA Wu R., Lin C.H., Coughlin S.R.;
 RT "Conserved structure and adjacent location of the thrombin receptor
 RT and protease-activated receptor 2 genes define a protease-activated
 RT receptor gene cluster.";
 RL Mol. Med. 2:349-357(1996).
 RN [3]
 RP SEQUENCE OF 1-74 FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=96372649; PubMed=8678993;
 RA Xue J., Jenkins N.A., Gilbert D.J., Copeland N.G., Sadler J.E.;
 RT "Structure and localization of the thrombin receptor gene on mouse
 RT chromosome 13.";
 RL Mamm. Genome 7:625-626(1996).
 CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; L03529; AAA040438.1;
 CC EMBL; U36757; AAB38308.1;
 CC EMBL; U36756; AAB38308.1; JOINED.
 CC EMBL; U55076; AAB00198.1;
 CC EMBL; U55075; AAB00198.1; JOINED.
 CC HSSP; P34996; 1DDO.
 CC MGD; MGI:101802; F2r.
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm_1; 1.

DR PROSITE; PS00237; G-PROTEIN RECEPTOR 1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR 1; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Blood coagulation.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 41 REMOVED FOR RECEPTOR ACTIVATION (BY
 FT SIMILARITY).
 FT CHAIN 42 430 PROTEINASE ACTIVATED RECEPTOR 1.
 FT DOMAIN 42 107 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 108 133 1 (POTENTIAL).
 FT DOMAIN 134 142 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 143 162 2 (POTENTIAL).
 FT DOMAIN 163 181 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 182 203 3 (POTENTIAL).
 FT DOMAIN 204 223 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 224 244 4 (POTENTIAL).
 FT DOMAIN 245 273 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 274 293 5 (POTENTIAL).
 FT DOMAIN 294 316 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 317 339 6 (POTENTIAL).
 FT DOMAIN 340 354 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 355 379 7 (POTENTIAL).
 FT DOMAIN 380 430 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 60 65 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 88 91 POLY-PRO.
 FT SITE 41 42 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 180 259 BY SIMILARITY.
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 162 162 F -> S (IN REF. 1).
 FT CONFLICT 189 189 R -> Y (IN REF. 1).
 FT CONFLICT 223 223 R -> G (IN REF. 1).
 FT CONFLICT 262 262 V -> L (IN REF. 1).
 FT CONFLICT 365 365 S -> T (IN REF. 1).
 SQ SEQUENCE 430 AA; 47790 MW; 395FD64FAE52C9BF CRC64;
 Query Match 25.3%; Score 540.5; DB 1; Length 430;
 Best Local Similarity 31.6%; Pred. No. 1.1e-23;
 Matches 124; Conservative 79; Mismatches 137; Indels 53; Gaps 13;
 QY 20 LILVAAG-----LLFLPVTVCSQGINVSDNSAKTLTIKSF--NGGPONTTEEPPLSDI 71
 Db 6 LLILVALGLSLGCLLSRVPMSPQESERTATVNP-----RSFFLRNPSENTFLVPLGDE 61
 QY 72 EG-----WTGATTITKAECPEDSISTLHVNNATIGVLRSSLSSTQVIPAIVYLLFVVG 123
 Db 62 EEEKNESVLLGKRAVYLNISLPHTPPPPPISEASGLTSPWLTLPMPSTVITFVVS 121
 QY 124 VPSNITVTLWKLRLTK-SISLVIFHTNLAIAADLLFCVTLPPKIAHYHLGNMNVFGEVMCR 182
 Db 122 LPLNLVLAIAVFLRMKKKPAVVMYMLHMAADVLSVLPFKISYVFSCTDMQFGSGMCR 181
 QY 183 ITTVYFNGMYCAIILITCMGINRVLAATHP---FTYOKLPKRSFSLMCGIWMVFLY 239
 Db 182 FATAFYNGMYNINMISIDRELAIVYPTQSLSWRTLGRANFT---CVVIWMAIMG 238
 QY 240 MLPFVILKQYHVLHSEITTHCHVDVACESPSPRFYFVSLAFPGFLIPFVIFCYTT 299
 Db 239 VVPLLEKQTRVPGNLTTHCHVDUSE-NLMQGFYSYFSAFSAFFLPLVPLVSTVCYTS 297
 QY 300 LTHKLK-----KDRILWGYIKAVLL---ILVIFTICFAPTNILVILHANYHY--- 345
 Db 298 IIRCLSSAVANRSKSR-----ALFLSAVFCIFVCGFTNVLLIVH---YLFLSD 347
 QY 346 -HNTDSLPFMYLIALCLGSLNSCLDPFLYFVMS 377
 Db 348 SPTGEAAYFAVLLCVSVSCCIDPLIYYVAS 380

RESULT 9
 ID PAR1_CR1LO STANDARD; PRT; 428 AA.

Q00991; Q60461;
 01-APR-1993 (Rel. 25, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
 GN F2R OR PAR1.
 OS Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetulus.
 NCBI_TaxID=10030;
 [1]
 SEQUENCE FROM N.A.
 RX TISSUE=Lung;
 RA MEDLINE=91348247; PubMed=1652467;
 RA Rasmussen U.B., Vouret-Craviari V., Jallat S., Schlesinger Y.,
 RA Pages G., Pavirani A., Lecocq J.P., Pouyssegur J.,
 RA Oberghen-Schilling E.;
 RT "cDNA cloning and expression of a hamster alpha-thrombin receptor
 RT coupled to Ca2+ mobilization.";
 RL FEBS Lett. 288:123-128(1991).
 RN [2]
 RP SEQUENCE OF 42-428 FROM N.A.
 RX MEDLINE=96028007; PubMed=7488069;
 RA Hartmann T., Grace M.B., Buzard G.S., Ruoss S.J.;
 RT "Thrombin receptor polymorphism in Chinese hamster";
 RL Biochem. Biophys. Res. Commun. 215:974-980(1995).
 CC -1- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X61958; CAA43957.1; -;
 CC EMBL; U34047; AAB6747.1; -;
 CC PIR; S17148; S17148.
 CC HSP; P34996; 1DDD.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm1; 1.
 CC PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 CC Blood coagulation; Polymorphism.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 41 REMOVED FOR RECEPTOR ACTIVATION (BY
 FT SIMILARITY).
 FT CHAIN 42 428
 FT DOMAIN 42 105
 FT TRANSMEM 106 131
 FT DOMAIN 132 140
 FT TRANSMEM 141 160
 FT DOMAIN 161 179
 FT TRANSMEM 180 201
 FT DOMAIN 202 221
 FT TRANSMEM 222 242
 FT DOMAIN 243 271
 FT TRANSMEM 272 291
 FT DOMAIN 292 314
 FT TRANSMEM 315 337
 FT DOMAIN 338 352
 FT TRANSMEM 353 377
 FT DOMAIN 378 428
 FT DOMAIN 60 63 ASP/GLU-RICH
 FT DOMAIN 301 304 POLY-SER.

FT SITE 41 42 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 178 257 BY SIMILARITY.
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 84 84 MISSING (IN AN ALLELE).
 FT CONFLICT 384 384 H -> T (IN REF. 1).
 SQ SEQUENCE 428 AA; 47602 MW; E5492AB2CD877E2F CRC64;
 Query Match 24.5%; Score 522.5; DB 1; Length 428;
 Best Local Similarity 32.7%; Pred. No. 1.1e-22;
 Matches 129; Conservative 80; Mismatches 128; Indels 57; Gaps 16;
 QY 20 LILVAAG-----LLFLPVTVCQSGINVSNSAKPTLTIKSF---NGGPQTEFEFFPLSD 70
 DB 6 LLLVAAGSLCGPLLSRVPVRQPESEMTDATVNP-----RSFFLRNPG-ENTFEILPLGD 60
 QY 71 IEGWTGATT-----TIKACPEDSDISLTHVNATIGYLRSSLSLSTQVIPAIYILLFVVG 123
 DB 61 EEEKNESTLPESGRAIYLNKSHSPAPLAFISEDAS-CYLTPSWLRLFIPSVYTFVVS 119
 QY 124 VPSNI--VTLMKLSLRTKISILVIFHTNLATDLFCVTLPFKIAYHLNGNNWVEGVMC 181
 DB 120 LPLNILAIAVFLKMKVKK-PAVYVYMLHMAVDFVSVLPKISYFSGDMQFGSGMC 178
 QY 182 RITTVFVYGNMYCAILITCMGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVWVFL 238
 DB 179 RFATAAFYCNMYASIMLMTVISIDRFILAVVPIQSLSWRTLGRANFT---CLVIWMAIM 235
 QY 239 YMLPEVLKQEHYLVHSEITTCCHVDVDAESPSPSFRFYFVSLAFPGFLIPFVILFCYT 298
 DB 236 GWVPLLKEQITVRPGLNITTCCHVDVNE-TLLQGFYSYFSAFSAVFLVPLIISTICVM 294
 QY 299 TLIHLKLS-----KDRIMLYIKAVLL---ILVITFCFAPNIIILVHHANYHYN 347
 DB 295 SIIRCLSSSVANRSKSR-----ALFLSAAFVCFVFCVPGPTNVLIMH---YLLLS 344
 QY 348 ---TDSLYFMYLIALCLGSLNSCLDPLFYFMS 377
 DB 345 DSPATEKAYFAYLLCVCSVSSCCIDPLIYYAS 378
 RESULT 10
 PAR1_RAT
 ID PAR1_RAT STANDARD; PRT; 432 AA.
 AC P26824;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
 GN F2R OR PAR1.
 OS Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 NCBI_TaxID=10116;
 [1]
 SEQUENCE FROM N.A.
 RP TISSUE=Aortic smooth muscle;
 RC MEDLINE=92381002; PubMed=1324917;
 RA Zhong C., Hayzer D.J., Corson M.A., Wick K., Runge M.S.;
 RT "Molecular cloning of the rat vascular smooth muscle thrombin
 RT receptor. Evidence for in vitro regulation by basic fibroblast growth
 RT factor.";
 RL J. Biol. Chem. 267:16975-16979(1992).
 CC -1- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M81642; AAA42274.1; -
 DR PIR; A43448; A43448.
 DR HSSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOPOSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Blood coagulation.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 45 REMOVED FOR RECEPTOR ACTIVATION (BY
 FT SIMILARITY).
 FT CHAIN 46 432 PROTEINASE ACTIVATED RECEPTOR 1.
 FT DOMAIN 46 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 135 1 (POTENTIAL).
 FT DOMAIN 136 144 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 145 184 2 (POTENTIAL).
 FT DOMAIN 165 183 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 184 205 3 (POTENTIAL).
 FT DOMAIN 206 225 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 226 246 4 (POTENTIAL).
 FT DOMAIN 247 275 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 276 295 5 (POTENTIAL).
 FT DOMAIN 296 318 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 319 341 6 (POTENTIAL).
 FT DOMAIN 342 357 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 358 381 7 (POTENTIAL).
 FT DOMAIN 382 432 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 64 68 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 87 93 POLY-PRO.
 FT SITE 45 46 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 182 261 BY SIMILARITY.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 432 AA; 48280 MW; DD032B97ABA4A605 C64;
 Query Match 24.0%; Score 512.5; DB 1; Length 432;
 Best Local Similarity 31.5%; Pred. No. 4.1e-22;
 Matches 126; Conservative 75; Mismatches 134; Indels 65; Gaps 13;
 QY 20 LILVAGLLF-----LPVTVQSGINVSNSAKPTLTIKSFNGGPONTPEERPLSD- 70
 DB 6 LLLVAVGLSLGGLLSSRVPMQPESEMYATPYNPRSPFLNPSDETEQPLGDE 65
 QY 71 -----IEGWTGATTTIKAECP-----EDSISTLVHNNATIGYLRSSLSQTVP 113
 DB 66 EKNESIPLEG--RAVYLNKSRFPMPPPFISED-----SGYLTSPWLTLP 113
 QY 114 AIYILLVGVGVSNIIVLKLKSLRTK-SISLVIFHTNLAIALLFCVTLFPKIAVHLGN 172
 DB 114 SVYTFVIVSLPLNLAIAVFFVRMKVKKPAVYMLHLMADVLVSVLPFKISYFSGT 173
 QY 173 NVVFGVWCRITTVVFGNMYCAILITCGINRYLATAPH---FTYQKLKRSFSLMC 229
 DB 174 DQWFGGMRFRFATAACYCMYASIMLMTVISIDRFLAVVPIQSLSWRLTGRANFT---C 230
 QY 230 GIWVNVFLVFPFVILKQYHLVHSEITTHDWDVACSPSPFRFYVSLAFFGFLP 289
 DB 231 VVIWMAINGVPLLEKEQTTPVGLNITTHDVLNE-TLLHGFYSYFSAFSAIFFLVP 289
 QY 290 FVLIIFCYTLTHLKLK-----KDRWLGVKAVLL---ILVFTICFAPTNIILVI 338
 DB 290 LIISTVCYTSIIRCLSSSAVANRSKSR-----ALFLSAAVFCIFVFCFGPTNLLIV 342

QY 339 HHANY-YHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 DB 343 HVLILLSDPGTETAYFAYLLCVCVTASVASCIDPLIYYVAS 382
 RESULT 11
 PARI PAPH
 ID PARI PAPH STANDARD; PRT; 425 AA.
 AC P56488;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 1 precursor (PARI-1) (Thrombin receptor).
 GN F2R OR PARI OR BTHR12.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Papio.
 OX NCBI_TaxID=9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shoji M., Hayzer D.J., Hanson S.R.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AF028727; AAB84191.1; -
 HSSP; P34996; 1DDD.
 InterPro; IPR000276; GPCR_Rhodpsn.
 Pfam; PF00001; 7tm.1; 1.
 PROSITE; PS00237; G-PROTEIN RECEPTOR 1.
 PROSITE; PS0262; G-PROTEIN RECEPTOR 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Blood coagulation.
 FT SIGNAL 1 26 POTENTIAL.
 FT PROPEP 27 41 REMOVED FOR RECEPTOR ACTIVATION.
 FT CHAIN 42 425 PROTEINASE ACTIVATED RECEPTOR 1.
 FT DOMAIN 42 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 128 1 (POTENTIAL).
 FT DOMAIN 129 137 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 138 157 2 (POTENTIAL).
 FT DOMAIN 158 176 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 177 198 3 (POTENTIAL).
 FT DOMAIN 199 218 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 219 239 4 (POTENTIAL).
 FT DOMAIN 240 268 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 269 288 5 (POTENTIAL).
 FT DOMAIN 289 311 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 312 334 6 (POTENTIAL).
 FT DOMAIN 335 350 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 351 374 7 (POTENTIAL).
 FT DOMAIN 375 425 CYTOPLASMIC (POTENTIAL).
 FT SITE 57 60 ASP/GLU-RICH (ACIDIC).
 FT SITE 41 42 CLEAVAGE (BY THROMBIN).
 FT DISULFID 175 254 BY SIMILARITY.
 FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to

Db 61 KNEISGLTEYRLVINKSPLOKQPAFISDASGYLTSSWLTFLVPSVYTGFWVWSLPLN 120
 QY 128 I--VTWKLSLRKTSLSLVFHNLAIDLFCVTLPKFIAYHLNGNNVFGVMCRIT 185
 Db 121 IMAIVFVLKWKVK--PAVVMUHLATADLVFVSLVPFKISYFSGDMQFSELCRFT 179
 QY 186 VVFGYNNYCAILLTCMGINRYLATAHP---FTYQKLPKRSFSLMGVWVWVFLYMLP 242
 Db 180 AAFYCNMYASILLMTWISIDFLAVVPMQSLSWRTLGRASFT--CLAIWALAIAGVVP 236
 QY 243 FVLKQEHVLVHSEITTCDDVDACSPSPFRFYFVSLAFGLPFLPFIIFCYTTLIH 302
 Db 237 LLKQEQITQVGLNITTCDDVLNETLLEGYIAYFSAFSAVF--FFVPLIISTVCYVSIIR 295
 QY 303 KLAS-----KDRILGWIKAVL--ILVFTICEAPTNIILVIHANYHH--NTD 349
 Db 296 CLSSAVANKSKSR-----ALFSAVFCIFICFGPTNVLIIAHY-SPLSHSTTE 347
 QY 350 SLVFMVLIALLGLSLNSCLDPFLYFVMS 377
 Db 348 AAFVAILLCVSVSSICCIDPLIYYAS 375

RESULT 14

PAR4 RAT
 ID PAR4 RAT STANDARD; PRT; 395 AA.
 AC Q920E0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-like 3)
 DE like 3) (Coagulation factor II receptor-like 3).
 GN F2RL3 OR PAR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Duoenum;
 RA Chien E.K., Marietti S., Mendoza J., Philippe M.;
 RT "Cloning of the rat protease activated receptor isoforms 3 and 4.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: Receptor for activated thrombin or trypsin coupled to G proteins that stimulate phosphoinositide hydrolysis. May play a role in platelets activation.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- PTM: A proteolytic cleavage generates a new amino terminus that functions as a tethered ligand.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF310216; AAL26790.1; -
 CC EMBL: AF269246; AAK58604.2; -
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC Pfam: PF00001; 7tm_1; 1
 CC PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
 CC PROSITE; PSS0262; G-PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Blood coagulation.

FT SIGNAL 1 16
 FT PROPEP 17 58
 FT
 FT CHAIN 59 395
 FT DOMAIN 59 93
 FT TRANSMEM 94 114
 FT DOMAIN 115 119
 FT TRANSMEM 120 140
 FT DOMAIN 141 161
 FT TRANSMEM 162 182
 FT DOMAIN 183 203
 FT TRANSMEM 204 224
 FT DOMAIN 225 254
 FT TRANSMEM 255 275
 FT DOMAIN 276 295
 FT TRANSMEM 296 316
 FT DOMAIN 317 330
 FT TRANSMEM 331 354
 FT DOMAIN 355 395
 FT SITE 58 59
 FT DOMAIN 98 101
 FT DISULFID 160 239
 FT CONFLICT 1 9
 FT CONFLICT 192 192
 SQ SEQUENCE 395 AA; 42943 MW; 43836D0DA22FAFAC CRC64;
 Query Match 23.2%; Score 496; DB 1; Length 395;
 Best Local Similarity 31.2%; Pred. No. 3.1e-21;
 Matches 104; Conservative 69; Mismatches 144; Indels 16; Gaps 4;
 QY 45 SAKPTLTITKSNFGGPGNTFEEPLSDIEGWTGATTIKAECPEDSISTLHVNNATIGYLR 104
 Db 39 SLRPTVELNESKSPKPNRGP-----GKPCANNSDLELPASSEALL-----LGV-- 85
 QY 105 SSLSTQVIPAIYLLFVGVGPSNIVTLMKLSLRKTSISLVIFHTNLAIADLCVTPLPK 164
 Db 86 --VPTRLVPAYIGLVVWVGLPANGALWVLAIVRPLPSTILLMNLAVADLLALVLP 143
 QY 165 IAYHLNGNNVFGVMCRITTVFVGNMYCAILLTCMGINRYLATAHPFTYQKLPKRSF 224
 Db 144 LVYHLRGQRWPEGEACRVATAALYGHMYGVSLLAAVSLDRYALVLSRARALRGQL 203
 QY 225 SLLMCGIVWVFLYMLPFVILKQEHVLVHSEITTCDDVDACSPSPFRFYFVSLAF 284
 Db 204 TTIILCVLWLSATLVLPVLFTHRQFLLAGSRMLCHDALPLAEQTSWHR--PAFICLAVL 262
 QY 285 GLIPFVFIIFCYTTLIHLKLSKDRIMWLGVIKAVLLILVIFTICFAPTNIIIVIHANY 344
 Db 263 GCFVPLLAMVLCYGATLALAAANGQRYSHAVRLTALVLFSAVAFAFTPSNVLVLVHYNPS 322
 QY 345 YINTSLSFMYLIALLGLSLNSCLDPFLYFVMS 377
 Db 323 PEAWGNLYGAYVYVPSLALSTLNSCVDPFIYYVS 355

RESULT 15

PAR4 MOUSE
 ID PAR4 MOUSE STANDARD; PRT; 396 AA.
 AC O88634;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-like 3) (Coagulation factor II receptor-like 3).
 GN F2RL3 OR PAR4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98389762; PubMed=9722561;

RESULT 18
 ID GP17_HUMAN STANDARD; PRT; 367 AA.
 AC Q13304; Q9UDZ6; O9UE21;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable P2Y purinoceptor GPR17 (P2Y-like receptor) (R12).
 GN GPR17.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=96145150; PubMed=8558062;
 RA Raport C.J., Schweickart V.L., Chantry D., Eddy R.L. Jr., Shows T.B.,
 RA Godiska R., Gray P.W.;
 RT "New members of the chemokine receptor gene family.";
 RT J. Leukoc. Biol. 59:18-23(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Hippocampus;
 RX MEDLINE=98181695; PubMed=9523551;
 RA Blaesus R.H., Weber R.G., Lichter P., Ogilvie A.;
 RT "A novel orphan G protein-coupled receptor primarily expressed in the
 RT brain is localized on human chromosomal band 2q21.";
 RL J. Neurochem. 70:1357-1365(1998).
 CC -!- FUNCTION: Putative receptor for purines coupled to G-proteins (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -!- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; U33447; AAB16746.1; -;
 CC EMBL; Y12546; CAA73144.1; -;
 CC EMBL; Z94154; CAB08107.1; -;
 CC EMBL; Z94155; CAB08108.1; -;
 CC HSSP; P34996; 1DDD.
 CC Genew; HGNC:4471; GPR17.
 CC MIM; 603071; -;
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHOOPS.
 CC PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
 CC PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Alternative splicing.
 CC
 CC DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 65 85 1 (POTENTIAL).
 CC DOMAIN 86 92 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 93 113 2 (POTENTIAL).
 CC DOMAIN 114 133 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 134 154 3 (POTENTIAL).
 CC DOMAIN 155 175 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 176 196 4 (POTENTIAL).
 CC DOMAIN 197 223 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 224 244 5 (POTENTIAL).
 CC DOMAIN 245 260 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 261 281 6 (POTENTIAL).
 CC DOMAIN 282 EXTRACELLULAR (POTENTIAL).
 CC

FT TRANSMEM 309 329 7 (POTENTIAL).
 FT DOMAIN 330 367 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 132 209 BY SIMILARITY.
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 1 28 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 367 AA; 40989 MW; 132FBE97BE83C60C CRC64;
 Query Match 18.7%; Score 398.5; DB 1; Length 367;
 Best Local Similarity 28.3%; Pred. No. 7.4e-16;
 Matches 99; Conservative 60; Mismatches 136; Indels 55; Gaps 9;
 QY 42 SNSAKPTLTITKSPNG---GPQMTPEEPPLSDIEGWTGATTITKACPEDSISTLHVNNA 98
 DB 23 SDSS-----QSMNGLEVAPPGLITNFSLATAE-----QCQGE----- 54
 QY 99 TIGYLRSLSSTOVIPAIYILLFVGVGPSNIVTLWKLRLTKS-LSLVIHTNLAIADLLF 157
 DB 55 -----TPLENNMLFASFYLLDFILALVGNLTALWLFIRDHKSGTPTANVFLMHLAVDLSC 108
 QY 158 CVTLPEKIAYLHNGNNWVFGEVMCRITTVVFYGNMYCAILILTCMGINRYLATAPFTYQ 217
 DB 109 VLVLPTRLVYHFGSNHWPFGEIACRLTGFLFLNMTASIFYLTCTISADRLAIVHPVKS 168
 QY 218 KLPKRSFSLLMCGIVVMVFLYMLPFVILKQYHLVHSEITTCDDVDVACESPSPSRFY 277
 DB 169 KLRRPLIYAHACAFLLVWVAVAPLLVSPQ-----TVQTNHTVVCLQLYREKASHHA 221
 QY 278 FVSLAPFGELIPFVLIIFCYTLIHLKLS-----KORIMGYKAVLLIIVITFCIPATN 333
 DB 222 LVSLA-VAFTFPFITVTCTYLLIIRSLRQGLRVEKRLTKAVRMTAIVLAIFLVCFVPVH 280
 QY 334 IILVIHNNYVYHNTDS-----LYFMYLIALGLSINSLCDPLFLYFVMSK 378
 DB 281 VNRSVTLVYHRSHGASCATORILANRITSCITSLNGALDPIIMFYFVAE 330

RESULT 19

CLT2_PIG
 ID CLT2_PIG STANDARD; PRT; 345 AA.
 AC Q9SN03;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteinyl leukotriene receptor 2 (CysLTR2).
 GN CYSLTR2 OR CYSLTR2.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
 RA Ohishi T., Soga T., Matsushime H., Furuichi K.;
 RT "Characterization of the cloned rat and porcine cysteinyl leukotriene
 RT receptors.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
 CC mediated via a G-protein that activates a phosphatidylinositol-
 CC calcium second messenger system (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AB052662; BAB60817.1; -;
 CC

```
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 44 64 1 (POTENTIAL).
FT DOMAIN 65 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 94 2 (POTENTIAL).
FT DOMAIN 95 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 145 3 (POTENTIAL).
FT DOMAIN 146 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 4 (POTENTIAL).
FT DOMAIN 176 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 226 5 (POTENTIAL).
FT DOMAIN 227 246 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOMAIN 268 287 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 288 308 7 (POTENTIAL).
FT DOMAIN 309 345 CYTOPLASMIC (POTENTIAL).
FT DISULFID 112 188 BY SIMILARITY.
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 345 AA; 5D1B1FB9BB95905 CRC64;

Query Match 18.5%; Score 395; DB 1; Length 345;
Best Local Similarity 29.9%; Pred. No. 1.1e-15;
Matches 97; Conservative 66; Mismatches 137; Indels 24; Gaps 10;

QY 86 PEDSISLTHVNATIGVLRSLSS-----TOVPAIYLLFVGVPSNIVTLWK-LSLR 137
DB 10 PSLSLSEME-PSNLTGNHNRSCTTENFKREYPIVILVIFWGLNGSGSIYVFLKPY 68

QY 138 TKSISLVIFHTNLAIADLLFCVTLPEKIAHYLNGNNWVFGVNMCRITTVVFGNMYCAIL 197
DB 69 KKSITSVNVFNLAIASDLLFTITLPRVDYLRAGNKFIDGTCRIMSYMYVNYSSIV 128

QY 198 ILTCMGINRYLATAHPTTYOKLPKRFSLLMCGIWMVFLYMLPFLVKQEHVHSEI 257
DB 129 FLTVLSVVRFLATVHPRLHTTSIRKANILCVGIWIFI---NASSTLKLKNGSEQKDNV 185

QY 258 TTCHDVVDACESPSFRFYFVSFLAFGLIPVILIFCVTLILHLKLS-----KDRILWG 313
DB 186 TLCELE--NSNKYTKLTKMYVALV-VGVLPFGTSLICVLLIIRALLKVEVPESGRLS 242

QY 314 YIKA---VLLILVIFTCFAPTNIILVIHANYHHN-TDSLVMYLIALLCLGSLNSCLD 369
DB 243 HRKALITVIALIIFLLCFLPYHVLRTLHLEWKADCKDRLHKAVAVTALAAANSCEF 302

QY 370 PFL-YFVMSKVVDQLNPXSMARP 392
DB 303 PFLYYFAGENFKDRLKSALRGRP 326

RESULT 20
P2YR_MELGA STANDARD; PRT; 362 AA.
AC P49652;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (GPI orphan receptor).
DE P2RY1.
GN Meleagris gallopavo (Common turkey).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
```



```

Db 122 VNLGSLFCTCSVHRYGTGVHPLKSLGRLKKNVYVSSLVWALVAVIAP-ILFYSG 180
QY 250 YHLVHSEITTCVCHVDACESPSPFRFYFVSLAFFGLIPFVIIIFCYTTTLHKLSKD- 308
Db 181 TGVRRKNTICYTT-ADEVLRYS-FVYSCTTVFNCIPFVILGCGLVKALYKDL 238
QY 309 -----RIWLGYIKAVLLILVITFCFAPTNIILVIH-HANYYYHN-----TDSLXF 353
Db 239 DNSPLRKRSYL-----VIILTVFAVSYPFPHVMTLNLRLARLDFQTPQMCAPFNKVA 293
QY 354 MYLIALCLGSLNSCLDPLFYFV 375
Db 294 TYQVTRGLASLNSCVDPILYFL 315

```

RESULT 21

```

P2Y8_XENLA
ID P2Y8_XENLA STANDARD; PRT; 537 AA.
AC P79928;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE P2Y purinoceptor 8 (P2Y8).
GN P2Y8.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Neural plate;
RX MEDLINE=97284734; PubMed=9139711;
RA Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;
RT "Early expression of a novel nucleotide receptor in the neural plate
of Xenopus embryos.";
RL J. Biol. Chem. 272:12583-12590 (1997).
CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, GTP AND ITP.
CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC MAY PLAY A KEY ROLE IN THE EARLY DEVELOPMENT OF NEURAL TISSUE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99953; CAA68213.1; .
DR HSPSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1. 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
FT FT 50 70 1 (POTENTIAL).
FT FT 71 79 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 80 100
FT FT 101 118 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 119 139
FT FT 140 161 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 162 182
FT FT 183 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 231
FT DOMAIN 232 254 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 255 275

```

```

FT DOMAIN 276 292 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 293 316 7 (POTENTIAL).
FT DOMAIN 317 537 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 193 BY SIMILARITY.
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 537 AA; 62024 MW; B2CF24812F3C19F2 CRC64;
Query Match 18.4%; Score 393; DB 1; Length 537;
Best Local Similarity 28.5%; Pred. No. 2e-15;
Matches 95; Conservative 62; Mismatches 116; Indels 60; Gaps 11;
QY 87 EDSISTLVNNATIGYLRSSL-----STQ-----VIPAIYILLFVGVPSNI 128
Db 3 EDIMATSYPTFLTTPYLPKMLNLTNDTDCVDFEGFKLLLPVYSVAVFVGLPLNI 62
QY 129 VTLWKLSLRTKSI-SLVIFHTNLATADLLFCVTLTPKIAHYHLNGNNMVGEVMCRITTV 187
Db 63 AAWWIFIAKQRPWNPTTVYMFNLALSDTLVLSLPTLVYVYADKNNWPFGEVLCKLVRFL 122
QY 188 FYGNMYCAIILTCMGINRYLATAHPFTYQKLPKRSFSLLMCGIVMWVFLVMLPFVILK 247
Db 123 FYANLYSSILFTCISVHRYGCVCHPITSLRRMNAKHAYVICALVWLSVTLCVLPNLI-- 180
QY 248 QEYHLVHSEI--TTCHDVVDACESPSPFRFY--YFVSLAFFGLIPFVIIIFCYTTLIHK 303
Db 181 --FVTVPKVKNTICHTDT---RPEDFARYVEYSTAIMCLLFGIPCLIIAGCYGLMTRE 234
QY 304 LKS-----KDRILGYIKAVLLILVITFCFAPTNIILVIHANYYY----- 345
Db 235 LMKPIVSGNQOTLPSYKKE---SIKTIIFVMAIFAICMPFHITRTL---YYARLLG 286
QY 346 ---HNTDSLYFMVLIALCLGSLNSCLDPLFYFV 375
Db 287 IKCYALNVINVTYKVTTRPLASANSCLDPILYFL 319
RESULT 22
P2Y9_HUMAN
ID P2Y9_HUMAN STANDARD; PRT; 370 AA.
AC Q99677; O15132;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2Y purinoceptor 9 (P2Y9) (Purine receptor 9) (G protein-coupled
receptor GPR23) (P2Y5-like receptor).
GN GPR23 OR P2Y9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=97225799; PubMed=9073069;
RA O'Dowd B.F., Nguyen T., Jung B.P., Marchese A., Cheng R.,
RA Heng H.H.Q., Kolakowski L.F. Jr., Lynch K.R., George S.R.;
RT "Cloning and chromosomal mapping of four putative novel human
G-protein-coupled receptor genes.";
RL Gene 187:75-81 (1997).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=97366605; PubMed=9223435;
RA Janssens R., Boeynaems J.M., Godart M., Communi D.;
RT "Cloning of a human heptahelical receptor closely related to the P2Y5
receptor.";
RL Biochem. Biophys. Res. Commun. 236:106-112 (1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: NOT DETECTED IN THE BRAIN REGIONS THALAMUS,
PUTAMEN, CAUDATE, FRONTAL CORTEX, PONS, HYPOTHALAMUS, HIPPOCAMPUS.

```

```
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U66578; AAC51301.1; -
DR EMBL; U90323; AAB62087.1; -
DR EMBL; U90322; AAB62088.1; -
DR EMBL; AF005419; AAB66322.1; -
DR HSSP; P34996; 1DDD.
DR Genew; HGNC:4478; GPR23.
DR MIM; 300086; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 44 64 1 (POTENTIAL).
FT DOMAIN 65 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 94 2 (POTENTIAL).
FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 133 3 (POTENTIAL).
FT DOMAIN 134 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176 4 (POTENTIAL).
FT DOMAIN 177 203 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 204 224 5 (POTENTIAL).
FT DOMAIN 225 254 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 255 275 6 (POTENTIAL).
FT DOMAIN 276 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 370 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 370 370 BY SIMILARITY.
FT DISULFID 111 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 192 192 F -> L (IN REF. 3).
SQ SEQUENCE 370 AA; 41895 MW; 2085752A3929E48 CRC64;
Query Match 18.1%; Score 387.5; DB 1; Length 370;
Best Local Similarity 29.3%; Pred. No. 3e-15;
Matches 93; Conservative 67; Mismatches 112; Indels 45; Gaps 11;
QY 87 EDSISLT--HVNNTATIG---YLRSSLSSTQVIPAIVYILLFVGVPSNIVTLKLSLRKSI 141
DB 12 QDSNSSLPRLGNATANTNCIVDDSPKYNLNGAVISVVFILGUTINSVLFVCFKMR 71
QY 142 S-LVIFHTNLAIADLLFCVTLPEKIAVHLNNGNWFGEVNCRIITTVFVGNMYCAILLT 200
DB 72 SETAIFITNLAVSDLLFVCTLPKIFNFN-RHWPFGDTLCKISGTAFLNIYVGMFLT 130
QY 201 CMGINRYLATAHPTFYQKLPKRFSLLMCGIWMVFLYMLPPIVLKQEVHLVHSITTC 260
DB 131 CISVDRLATVIFPFRSRTITRNSAIVCAGWILVLSGGISASLF-----STTNV 181
QY 261 HDVVDADESSSPREF--YFVSIAFF---GFLIPFVILIFCVTTLIHLK----- 305
DB 182 NNATTCFEGFSKRWKTYLSKTIFFIEVVGFIPIILNVCSSVLTLRKATLSQIG 241
QY 306 -SKDRILWGIKAVLLILVIFTICFAPTNILVIHH-----ANYYYHNTDSLYPMYLI 357
DB 242 TNKKV---LKMIVHMAVFCVFPYNSVFLYALVRSQAITNCFLERFAXI--MYPI 295
QY 358 ALCLGSLNSCLDPFLYF 374
DB 296 TLLCATLNCDFPIFY 312
```

```
RESULT 23
P2Y5 CHICK
ID P2Y5 CHICK STANDARD; PRT; 308 AA.
AC P32250;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE P2Y purinoceptor 5 (P2Y5) (Purinergic receptor 5) (6H1).
GN P2RY5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=93329058; PubMed=8393036;
RA Kaplan M.H., Smith D.I., Sundick R.S.;
RT "Identification of a G protein coupled receptor induced in activated
RT T cells.";
RL J. Immunol. 151:628-636(1993).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: INDUCED IN ACTIVATED T-CELLS.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06109; AAB06587.1; -
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW Palmitate.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 17 43 1 (POTENTIAL).
FT DOMAIN 44 52 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 53 76 2 (POTENTIAL).
FT DOMAIN 77 89 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 90 109 3 (POTENTIAL).
FT DOMAIN 110 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 151 4 (POTENTIAL).
FT DOMAIN 152 178 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 179 206 5 (POTENTIAL).
FT DOMAIN 207 224 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 225 250 6 (POTENTIAL).
FT DOMAIN 251 269 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 270 289 7 (POTENTIAL).
FT DOMAIN 290 308 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 281 281 PALMITATE (BY SIMILARITY).
FT DISULFID 86 165 POTENTIAL.
SQ SEQUENCE 308 AA; 35597 MW; 4214E969633B6F7D CRC64;
Query Match 17.9%; Score 383; DB 1; Length 308;
Best Local Similarity 30.3%; Pred. No. 4.6e-15;
Matches 90; Conservative 59; Mismatches 104; Indels 44; Gaps 9;
QY 115 IYVILFVGVPSNIVTLW----KLSLRKTSISLVIFHTNLAIADLLFCVTLPEKIAHNL 170
DB 20 VFSMVVFLGLIANCAVIAIFTTLKVRNNTTYML---NLASDLLFVTLFRI--YFV 75
QY 171 GNNVFGVEMCRITTVFVGNMYCAILLTTCGINRYLATAHPTFYQKLPKRFSLLMCG 230
```

```
Db 76 VRNPFQDVLCKISVTLFYTNMVGSLFLTCISVDRLAIVHPFRSKTLRTKRNARIVCV 135
QY 231 IVMVMVFLYMLPFVILKQEHVHLVHSEITTCIDVDVADACESPSFRFFYFVSLAFF-----GF 286
Db 136 AVMITVLGASTPASFOSTNRQNTQRTCFE-----NPESTWTKYLSRIVFIEIVGF 190
QY 287 LIPFVILIIFFYTTLIHLK-----SKRIMWGYIKAVLLIIVITICFAPTNIL 336
Db 191 FIPILINVTCSWVLRNLKPLSLRNKLSKKV-----LKMIFVHLVIFCFECPVPYNITL 246
QY 337 VIHHANVYVHNTD-----SLYFMYLIALCLGSLNSCLDPFLYFVMSKVVDQLN 384
Db 247 IL-----YSLMRTQWNCVSWTAIRTMYPVTLCAVNSCCFDEIVYFTSDTNSELD 299

RESULT 24
CLUT2_RAT
ID CLUT2_RAT STANDARD; PRT; 309 AA.
AC Q924T9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 2 (CysLTR2) (RSBPT32).
GN CysLTR2 OR CysLT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT "Characterization of the cloned rat and porcine cysteinyl leukotriene
RL receptors."
CC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB052661; BAB60816.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN_RECPT_F1_1; FALSE_NEG.
DR PROSITE; PSS0262; G-PROTEIN_RECPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 27 47 1 (POTENTIAL).
FT DOMAIN 48 56 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 57 77 2 (POTENTIAL).
FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 119 3 (POTENTIAL).
FT DOMAIN 120 137 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 138 158 4 (POTENTIAL).
FT DOMAIN 159 187 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 208 5 (POTENTIAL).
FT DOMAIN 209 229 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT DOMAIN 251 271 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 272 292 7 (POTENTIAL).
FT DOMAIN 293 309 CYTOPLASMIC (POTENTIAL).
FT DISULFID 95 171 BY SIMILARITY.
FT CARBOHYD 14 14 N-LINKED (GLCNAC...) (POTENTIAL).
FT 166 N-LINKED (GLCNAC...) (POTENTIAL).
FT 166
```

```
FT CARBOHYD 167 167 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 309 AA; 35507 MW; A122AC8177879D56 CRC64;
Query Match 17.9%; Score 382.5; DB 1; Length 309;
Best Local Similarity 31.4%; Pred. No. 5e-15;
Matches 93; Conservative 53; Mismatches 113; Indels 37; Gaps 8;
QY 97 NATIGYLRSSLSGTQVPAIYIILLFVGVPSNIVTLWK-LSLRKTSISLVFHTNLAIDL 155
Db 14 NCTI-----ENFKRDFYPIIYLIIFVWGALNGFISVVFQTKYSTVNVFVNLNLAISDF 69
QY 156 LFCVTLPEKIAHNGNNWVFEVNCRTITVVVFGNYMCAIILITCMGINRVLATAHPT 215
Db 70 LFISTLPFRADYNFRGSDWIFGDWACRIMSLSYNNMTSYIYFLTVLSVRLATAHPTQ 129
QY 216 YQKLPRKRSFLMCGIWMVFLYMLPFVILKQEHVHSE-----ITTCIDVDVADCESP 270
Db 130 MLIHTSVRSANILCGIHW-----FIMASSGLLLKHGQKKNTLTCFEL----- 174
QY 271 SSFRFYFVSLAFF---GFLIPFVILIIFCYTTLIHLKSKDRIMLG-----YIKAVL 319
Db 175 NLQKFNILVILNYIALGVGFLLPFFILATCYLLIIRVLLKVEIPESGPRDAQRKALTTIV 234
QY 320 LILVIFTICFAPTNILVIHHANYHHN-TDSLYFMYLIALCLGSLNSCLDPFLYF 374
Db 235 IAMIIFLLCFLPYHALRTIHLVTDADSCMDLHKATVITLTAANSFCNPFLLY 290

RESULT 25
CLUT2_HUMAN
ID CLUT2_HUMAN STANDARD; PRT; 346 AA.
AC Q9NS75; Q9HCQ2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 2 (CysLTR2) (PSEC0146) (HG57) (HPN321).
GN CysLTR2 OR CysLT2 OR CysLT2R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20374466; PubMed=10913337;
RA Takasaki J., Kamohara M., Matsumoto M., Saito T., Sugimoto T., Ota T.,
RA Nishikawa T., Kawai Y., Masuho Y., Isogai T., Suzuki Y., Sugano S.,
RA Furuichi K.;
RT "The molecular characterization and tissue distribution of the human
RL cysteinyl leukotriene CysLTR2 receptor."
RL Biochem. Biophys. Res. Commun. 274:316-322(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20459128; PubMed=10851239;
RA Heise C.E., O'Dowd B.F., Figueroa D.J., Sawyer N., Nguyen T.,
RA Im D.-S., Stocco R., Bellefeuille J.N., Abramovitz M., Cheng R.,
RA Williams D.L. Jr., Zeng Z., Liu Q., Ma L., Clements M.K., Coulombe N.,
RA Liu Y., Austin C.P., George S.R., O'Neill G.P., Metters K.M.,
RA Lynch K.R., Evans J.F.;
RT "Characterization of the human cysteinyl leukotriene 2 receptor."
RL J. Biol. Chem. 275:30531-30536(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20545741; PubMed=11093801;
RA Nothacker H.-P., Wang Z., Zhu Y., Reinscheid R.K., Lin S.H.S.,
RA Civelli O.;
RT "Molecular cloning and characterization of a second human cysteinyl
RL leukotriene receptor: discovery of a subtype selective agonist."
RN Mol. Pharmacol. 58:1601-1608(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Dunn M.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
```

```

RN [5]
RP SEQUENCE OF 17-346 FROM N.A.
RA Suga H.;
RT "Homo sapiens cysteinyl leukotriene receptor 1 like receptor.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system. Stimulation by BAY u9773, a
CC partial agonist, induces specific contractions of pulmonary veins
CC and might also have an indirect role in the relaxation of the
CC pulmonary vascular endothelium. The rank order of affinities for
CC the leukotrienes is LTC4 > LTD4 > LTE4.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in the
CC heart, placenta, spleen, peripheral blood leukocytes and adrenal
CC gland. In lung, expressed in the interstitial macrophages, and
CC slightly in smooth muscle cells.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB039269; BAB03601.1; -
DR EMBL; AF254664; AAK17281.1; -
DR EMBL; AF279611; AAK69485.1; -
DR EMBL; AL137118; CAC29102.1; -
DR EMBL; AB041644; BAB16379.1; -
DR MIM; 605666; -
DR InterPro; IPR004071; Cysleuk_receptor.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01533; CYSLTRRCPTR.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; FALSE_NEG.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_2; 1.
DR PROSITE; PS50262; G_PROTEIN_RECP_TM; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 42 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 43 62 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 64 72 1 (POTENTIAL).
FT TRANSMEM 73 93 2 (POTENTIAL).
FT DOMAIN 94 123 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 124 144 3 (POTENTIAL).
FT DOMAIN 145 153 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 154 174 4 (POTENTIAL).
FT DOMAIN 175 204 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 205 225 5 (POTENTIAL).
FT DOMAIN 226 245 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 246 266 6 (POTENTIAL).
FT DOMAIN 267 286 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 287 307 7 (POTENTIAL).
FT DOMAIN 308 346 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 187 BY SIMILARITY.
FT CARBOHYD 20 20 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 346 AA; 39635 MW; EB54A4A2DDCE5EE4 CRC64;

Query Match 17.78; Score 378.5; DB 1; Length 346;
Best Local Similarity 28.98; Pred. No. 9.1e-15;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

QY 86 PEDSISLTHVNNATIGYRLSLSQVIPAIVYLLVGVPSNIVTLWK-LSLRTKGISLV 144
DB 19 PNGTFNNSENCTI-----ENFKREPPPIVLIIFPVGVLGSLIVFLQPKKTSVN 74
QY 145 IFTNLAIALDLFCVTLFPFKIAYHLNGNNVWGEVWCRIITVVYFGNYCAILILCMGI 204
: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:

```

RESULT 26

P2Y2 MOUSE

ID P2Y2 MOUSE STANDARD; PRT; 373 AA.

AC P35383; O09031; QSCPZ4;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)

DE (Purinergic receptor).

DE P2RY2 OR P2RU1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]_TaxID=10090;

RN SEQUENCE FROM N.A.

RP MEDLINE=93281707; PubMed=7685114;

RX Lustig K.D., Shiau A.K., Brake A.J., Julius D.;

RT "Expression cloning of an ATP receptor from mouse neuroblastoma

RT cells";

RL Proc. Natl. Acad. Sci. U.S.A. 90:5113-5117(1993).

RN [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=96316177; PubMed=8739319;

RA Enomoto K., Furuya K., Moore R.C., Yamagishi S., Oka T., Maeno T.;

RT "Expression cloning and signal transduction pathway of P2U receptor

RT in mammary tumor cells";

RL Biol. Signals 5:9-21(1996).

RN [3]

RN SEQUENCE FROM N.A.

RX STRAIN=C57BL/6J; TISSUE=Head, and Liver;

RX MEDLINE=21085560; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Scorch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [4]

RN SEQUENCE FROM N.A.

RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP MUTAGENESIS OF LYS-107; ARG-110; HIS-262; ARG-265; LYS-289 AND
RP ARG-292
RX MEDLINE=95181393; PubMed=7876172;
RA Erb L., Garrad R.C., Wang Y., Quinn T., Turner J.T., Weisman G.A.;
RT "Site-directed mutagenesis of P2U purinocceptors. Positively charged
RT amino acids in transmembrane helices 6 and 7 affect agonist potency
RT and specificity.";
RT J. Biol. Chem. 270:4185-4188 (1995).
CC -1- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC activate a phosphatidylinositol-calcium second messenger system.
CC The affinity range is UTP > ATP > ATP-gamma-S >> 2-methylthio-ATP
CC = ADP.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND
CC BRAIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14751; AAA39871.1; -;
DR EMBL; S83099; AAB50735.1; -;
DR EMBL; AK005013; BAB23746.1; -;
DR EMBL; AK017378; BAB30719.1; -;
DR EMBL; BC006613; AAH06613.1; -;
DR HSP; P34996; 1DD0.
DR MGD; MGI:105107; P2ry2.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOPOPS.
DR PROSITE; PS00237; G PROTEIN RECP Fl_1; 1.
DR PROSITE; PS00262; G PROTEIN RECP Fl_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 33 59 1 (POTENTIAL).
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 71 93 2 (POTENTIAL).
FT DOMAIN 94 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 129 3 (POTENTIAL).
FT DOMAIN 130 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 172 4 (POTENTIAL).
FT DOMAIN 173 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 220 5 (POTENTIAL).
FT DOMAIN 221 246 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 247 269 6 (POTENTIAL).
FT DOMAIN 270 287 7 (POTENTIAL).
FT TRANSMEM 288 309 7 (POTENTIAL).
FT DOMAIN 310 373 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 106 183 BY SIMILARITY.
FT MUTAGEN 107 107 K->I: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 110 110 R->L: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 262 262 H->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 265 265 R->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 289 289 K->I: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 292 292 R->L: DECREASE IN RECEPTOR ACTIVATION.
FT CONFLICT 17 17 E->D (IN REF. 2).
FT CONFLICT 120 120 S->R (IN REF. 2).
FT CONFLICT 125 125 T->N (IN REF. 2).
FT CONFLICT 196 196 V->M (IN REF. 2).
FT CONFLICT 263 263 V->L (IN REF. 3).
FT CONFLICT 355 355 D->N (IN REF. 2).
FT CONFLICT 369 369 KDI -> PYV (IN REF. 2).

SQ SEQUENCE 373 AA; 42174 MW; 590BBE502E41B3AE CRC64;
Query Match 17.7%; Score 378.5; DB 1; Length 373;
Best Local Similarity 26.9%; Pred. No. 9.6e-15;
Matches 100; Conservative 67; Mismatches 144; Indels 61; Gaps 13;
QY 69 SDIEGWTGATTIKAECPEDSISTLHVNNATGYLRSSLSSTQVIPAIIYILLFVVGVPVSI 128
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 52
3 ADLEPW--NSTINGTWEGDELGYKCRFNEDPKYV-----LLPVSYGVVGVGLCLNV
QY 129 VTLMKLSLRTK---SISLVIPHTNLAIADLLFCVTLPFKIAVHLGNWVFGVMCRIT 185
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 110
53 VALYIFLCKLKTWNASTYMPH--LAVSDLSAAASLPLLVYYYARGDHPFSTVLCKLVR
QY 186 VVFGNMYCAIILATCMGINRYLATAHPFTYOKLPKRSFSLMCGVWVWVFLYMLP--- 242
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 170
111 FLFYTNLYCSILFLACISVHRCLGVLRPLSLRWGRARVAAVAVWVLVLAQAPVLY
QY 243 FVLKQEVHLVHSEITTHCHVDVADACESPSSPRFYFVSLAPFGFLIPFVILFYCYTTLH 302
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 223
171 FVTTTS-----VRGTRITCHD-TSARELFSHFVAYSSVLMGLL-FAVPFSVILVYVLMAR
QY 303 KL-----KSKDRIMLVGVIKAVLLILVIFTICFAPTNILVIHHA----- 344
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 277
224 RLKPAYGTGTGLPRAKRKS-----VRTIALVAVFALCFLPFHVTRTLYYSFRSLDLS
QY 345 YHNTDSLVFMYLIALCLGSLNSCLDPFLYFVMSKV-----DOLNPXSAMARP----- 392
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 337
278 CHTLNAINWAYKITPLASANSCLDPVLYFLAGQLRVFARDAKPTETPTSPQARRKLG
QY 393 LXPRRRDIWEDI 404
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 338 LHRPNRTVRKDL 349
RESULT 27
CLT1 HUMAN STANDARD; PRT; 337 AA.
ID CLT1 HUMAN AC Q9Y271;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 1 (CysLTR1) (Cysteinyln leukotriene D4
DE receptor) (LTD4 receptor) (HG55) (HMTMF81).
GN CYSLTR1 OR CYSLTR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=99318129; PubMed=10391245;
RA Lynch K.R., O'Neill G.P., Liu Q., Im D.-S., Sawyer N., Metters K.M.,
RA Coulombe N., Abramovitz M., Figueroa D.J., Zeng Z., Connolly B.M.,
RA Bai C., Austin C.P., Chateaufort A., Stocco R., Greig G.M.,
RA Kargman S., Hooks S.B., Hosfield E., Williams D.L. Jr.,
RA Ford-Hutchinson A.W., Caskey C.T., Evans J.F.;
RT "Characterization of the human cysteinyl leukotriene CysLTR1
RT receptor.";
RL Nature 399:789-793 (1999).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood monocytes, Leukocyte, and Spleen;
RX MEDLINE=99393629; PubMed=10462554;
RA Sarau H.M., Ames R.S., Chambers J., Ellis C., Elshourbagy N.,
RA Foley J.J., Schmidt D.B., Muccitelli R.M., Jenkins O., Murdock P.R.,
RA Herxity N.C., Halsey W., Sathe G., Muir A.I., Nuthulaganti P.,
RA Dytko G.M., Buckley P.T., Wilson S., Bergsma D.J., Hay D.W.P.;
RT "Identification, molecular cloning, and characterization
RT of a cysteinyl leukotriene receptor.";
RL Mol. Pharmacol. 56:657-663 (1999).
CC -1- FUNCTION: Receptor for cysteinyl leukotrienes mediating

RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96257237; PubMed=8666290;
 RA Leon C., Vial C., Cazenave J.-P., Gachet C.;
 RT "Cloning and sequencing of a human cDNA encoding endothelial P2Y1
 RL purinoceptor.";
 RL Gene 171:295-297 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96158962; PubMed=8579591;
 RA Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,
 RA Kunapuli S.P.;
 RT "Cloning and chromosomal localization of the human P2Y1
 RL purinoceptor.";
 RL Biochem. Biophys. Res. Commun. 218:783-788 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96205320; PubMed=8630005;
 RA Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
 RA Boeynaems J.M.;
 RT "Cloning and tissue distribution of the human P2Y1 receptor.";
 RL Biochem. Biophys. Res. Commun. 221:588-593 (1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RL Leon C., Vial C., Weber J., Cazenave J.-P., Gacher C.;
 RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
 RC TISSUE=Platelet;
 RX MEDLINE=98113162; PubMed=9442040;
 RA Jin J., Daniel J.L., Kunapuli S.P.;
 RT "Molecular basis for ADP-induced platelet activation. II. The P2Y1
 RL receptor mediates ADP-induced intracellular calcium mobilization and
 RT shape change in platelets.";
 RL J. Biol. Chem. 273:2030-2034 (1998).
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 CC ATP AND ADP. IN PLATELETS BINDING TO ADP LEADS TO MOBILIZATION OF
 CC INTRACELLULAR CALCIUM IONS VIA ACTIVATION OF PHOSPHOLIPASE C. A
 CC CHANGE IN PLATELET SHAPE, AND PROBABLY TO PLATELET AGGREGATION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- INDUCTION: REPPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS
 CC A3PSPS, A3PSP AND A2PSP. THESE INHIBIT CALCIUM ION MOBILIZATION
 CC AND SHAPE CHANGE IN PLATELETS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z49205; CAA89066.1; -
 CC EMBL; U42030; AAA97873.1; -
 CC EMBL; U42029; AAA97872.1; -
 CC EMBL; S81950; AAB47091.1; -
 CC EMBL; AJ006945; CAA07339.1; -
 CC EMBL; AF018284; AAB94556.1; -
 CC HSPSP; P34996; 1DDD.
 CC Genew; HGNC:8539; P2RY1.
 CC MIM; 601167; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHOOPS.
 CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
 KW Blood coagulation.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 74 1 (POTENTIAL).
 FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 188 4 (POTENTIAL).
 FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 219 238 5 (POTENTIAL).
 FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 266 285 6 (POTENTIAL).
 FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 304 328 7 (POTENTIAL).
 FT DOMAIN 329 373 BY SIMILARITY.
 FT DISULFID 124 202 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 138 138 MISSING (IN REF. 1).
 SQ SEQUENCE 373 AA; 42071 MW; 4DC7C668B4145392 CRC64;
 Query Match 17.6%; Score 376.5; DB 1; Length 373;
 Best Local Similarity 27.0%; Pred. No. 1.2e-14;
 Matches 95; Conservative 72; Mismatches 130; Indels 55; Gaps 12;
 QY 46 AKPTLTIKSFNGGQNTFEFPPLSDIEGWTGATTTIKAECPEDSISTLHVNA-TIGYLK 104
 Db 8 AVPNGTDAFLAGPGSS-----WGNSTVA-----STRAVSSSFKALTK 46
 QY 105 SLSLSTQVTPAIYLLFVVGVPNSIVTLWKLSTK---SISLVIFHTNLAIALDLPCVTL 161
 Db 47 TGFQFYYPVAVYILVFIIGLGNVAIWMFVPHMKPWSGISVYMF--NLALADFLYVLT 104
 QY 162 PKFIAYHLNGNNVFGVWVCRTTIVFYGNMYCAILLTCMGINRYLATAHPTQKLPK 221
 Db 105 PALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFTICSAHRYSGVGVPLKSLGRUK 164
 QY 222 RSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCCHVDVDACESSPFRFYFVSL 281
 Db 165 KQNAICISVLWLVVVAISP-ILFVSGTGVKKNITCYDT---SDEYLSRFYIYS 219
 QY 282 --AFTGFLIPFVILFCYTTLIHKLSKD-----RIWLGYKAVULLIVITFCFA 330
 Db 220 CTTVAMFCVPLVLILGCGYGLVIRALYIKDLNSPLRRKSIYL-----VIIIVTFAVSYI 274
 QY 331 PTNIIIVH-HANYVYHN-----TDSLYFMVLIACLGSLNSCLDPLFYFV 375
 Db 275 PFHVMTNLRARLDFOPTAMCAFNDRVYATVQVTRGLASLNSCVDPIILFL 326
 RESULT 31
 P2Y4_MOUSE STANDARD; PRT; 361 AA.
 ID P2Y4_MOUSE AC Q9JUS7;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoceptor 4 (P2Y4).
 GN P2RY4 OR P2Y4R.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=21185993; PubMed=11290369;
 RA Suarez-Huerta N., Pouillon V., Boeynaems J.-M., Robaye B.;
 RT "Molecular cloning and characterization of the mouse P2Y4 nucleotide
 RL receptor.";
 RL Eur. J. Pharmacol. 416:197-202 (2001).
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

S Bos taurus (Bovine).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
C Bovidae; Bovinae; Bos.
X NCBI_TaxID=9913;
N [1]

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement) or send an email to license@isb-sib.ch).

EMBL; X87628; CAA60958.1; .

1	DOMAIN	1	EXTRACELLULAR (POTENTIAL) .
2	TRANSMEM	52	1 (POTENTIAL) .
3	TRANSMEM	53	CYTOPLASMIC (POTENTIAL) .
4	DOMAIN	75	2 (POTENTIAL) .
5	TRANSMEM	88	EXTRACELLULAR (POTENTIAL) .
6	DOMAIN	110	3 (POTENTIAL) .
7	TRANSMEM	127	CYTOPLASMIC (POTENTIAL) .
8	DOMAIN	148	4 (POTENTIAL) .
9	TRANSMEM	167	EXTRACELLULAR (POTENTIAL) .
10	DOMAIN	189	5 (POTENTIAL) .
11	TRANSMEM	219	CYTOPLASMIC (POTENTIAL) .
12	DOMAIN	239	6 (POTENTIAL) .
13	TRANSMEM	266	EXTRACELLULAR (POTENTIAL) .
14	DOMAIN	286	7 (POTENTIAL) .
15	TRANSMEM	304	CYTOPLASMIC (POTENTIAL) .
16	DOMAIN	329	BY SIMILARITY .
17	TRANSMEM	373	N-LINKED (GLCNAC. . .) (POTENTIAL) .
18	DISULFID	124	N-LINKED (GLCNAC. . .) (POTENTIAL) .
19	CARBOHYD	11	N-LINKED (GLCNAC. . .) (POTENTIAL) .
20	CARBOHYD	27	N-LINKED (GLCNAC. . .) (POTENTIAL) .
21	CARBOHYD	113	N-LINKED (GLCNAC. . .) (POTENTIAL) .
22	CARBOHYD	197	N-LINKED (GLCNAC. . .) (POTENTIAL) .

Query Match 17.4%; Score 372; DB 1; Length 373;
Best Local Similarity 27.8%; Pred No. 2.2e-14;
Matches 94; Conservative 63; Mismatches 139; Indels 42; Gaps 10;

QY 59 PONTFEFFPLSDIEGTGATTIKARCPEDSISTLHVNATIGYLRSSLSSTQVIPAIIYL 118
 DB 10 PNCITDTAFLADPSPGNSTVSTAASVAFKALTKTGQFYLL-----PAVYIL 60
 QY 119 LFVVGVPNSIVTLWKLRLTK-----SISLVIFHTNLAIDLFCVTLPPFKIAYHLNGNNVV 175
 DB 61 VFIIGFLGNSVAIWMFVFHMKPSGISVYMF--NLALADFLVTLTLPALIFYFNKTDWI 118
 QY 176 FGEVMCRITTVYGNMYCAILLTTCGNYRLATAHPTTYQKLPKRSLSLMCGIWMV 235
 DB 119 FGDAMCKLQRFIFHVNLYGSIILCTISAHRYSGVVYPLKSLGRKKNAVVISLVWLI 178
 QY 236 VFVLMPLFVILKQYHLVHSEITTCDDVDADSPSSFRFYFVSVL--AFFGFLIPEVII 293
 DB 179 VVVGISP-ILFYSGTGIRKNTTCTCYDTT-----SDYLSRYFYISCTTVAFECVPLVI 233
 QY 294 IFCYTTLIHLKSKD-----RIWLGYIKAVLLILVIFTCFAPTNTIILVIH-HANY 343
 DB 234 LGCYGLIVRALIYKLDNSPLRKSIVL-----VIVLTVFAVSYPFHVMTKMLRL 288
 QY 344 YHNN-----TDSLFWYLIALLCLGSLNSCLDPLFYFV 375
 DB 289 DFQTPENCAFNDRVYATYQVTRGLASLNSCVDPLIYFL 326
 RESULT 33
 ID GP34_HUMAN STANDARD; PRT; 381 AA.
 AC Q9UPC5; O95853;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable G protein-coupled receptor GPR34.
 GN GPR34.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=99326137; PubMed=10395919;
 RA Schoneberg T., Schulz A., Grosse R., Schade R., Henklein P.,
 RA Schultz G., Gudermann T.;
 RT "A novel subgroup of class I G-protein-coupled receptors.";
 RL Blochim. Biophys. Acta 1446:57-70(1999).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99156852; PubMed=10036181;
 RA Marchese A., Sawzdargo M., Nguyen T., Cheng R., Heng H.H., Nowak T.,
 RA Im D.-S., Lynch K.R., George S.R., O'Dowd B.F.;
 RT "Discovery of three novel orphan G-protein-coupled receptors.";
 RL Genomics 56:12-21(1999).
 [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20434921; PubMed=10982042;
 RA Jacob J.K., Broghammer M., Pesch K., Zrenner E., Berger W.,
 RA Meindl A., Fusch C.M.;
 RT "Physical mapping and exclusion of GPR34 as the causative gene for
 RT congenital stationary night blindness type 1.";
 RL Hum. Genet. 107:89-91(2000).
 [4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Placenta;
 RA Isegal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Iehi S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 [5]
 RN SEQUENCE FROM N.A.
 RP

RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ORPHAN RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: BROADLY EXPRESSED.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF039686; AAD50531.1; -;
 DR EMBL; AF118670; AAD17248.1; -;
 DR EMBL; AK027780; BAB5362.1; -;
 DR EMBL; BC020678; AAB30678.1; -;
 DR Genew; HGNC:4490; GPR34.
 DR MIM; 300241; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL_2; 1.
 KW Receptor; G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 61
 FT TRANSMEM 62 82
 FT DOMAIN 83 88
 FT TRANSMEM 89 109
 FT DOMAIN 110 128
 FT TRANSMEM 129 149
 FT DOMAIN 150 171
 FT TRANSMEM 172 192
 FT DOMAIN 193 216
 FT TRANSMEM 217 237
 FT DOMAIN 238 269
 FT TRANSMEM 270 290
 FT DOMAIN 291 310
 FT TRANSMEM 311 331
 FT DOMAIN 332 381
 FT DISULFID 127 204
 FT CARBOHYD 28 28
 FT CARBOHYD 36 36
 FT CARBOHYD 42 42
 FT CARBOHYD 200 200
 FT CARBOHYD 295 295
 FT CONFLICT 181 181
 FT CONFLICT L -> V (IN REF. 1).
 SQ SEQUENCE 381 AA; 43860 MW; 491FC01685624379 CRC64;
 Query Match 17.2%; Score 367; DB 1; Length 381;
 Best Local Similarity 28.3%; Pred. No. 4.3e-14;
 Matches 105; Conservative 64; Mismatches 130; Indels 72; Gaps 14;
 QY 49 TLTIKSFNGGPQNT-----FEFFPLSDIEGTGATTIKARCPEDSISTLHVNATI 100
 DB 7 TMTTTSVSWPYSSSHRMFRITNHSDDPPQNFSAFNPVT-----CPMD----- 49
 QY 101 GYLRSLSLSTQVIPAIIYLLFVVGVPNSIVTLWKL-LSLRKTSISLVIFHTNLAIDLFCV 159
 DB 50 ----EKLJSTVLTTSYSVIFVGLVGNIIALVFLGIHKKRNSIQIYLLNVAIDLILF 105
 QY 160 TLPPFKIAYHLNGNNVFGVEVMCRITTVYFVGNMYCAILLTTCGNYRLATAHPTTYQKL 219
 DB 106 CLPFRIMYHINQNTKTLGVILCKVYVGTFLFYNNMYSIILLGFLISLDRIYKINRSIQORKA 165
 QY 220 PKRSFSLMCGIWMVWV---FLYMLPFVILKQYHLVHSEITTCDDVDADSPSSFRFY 276
 DB 166 ITTKOSIYVCCIVWMLALGGFUTMI-ILTLKKGGH-----NSTMCFHYRDKHNAKGAIFN 220
 QY 277 YFVSLAFFGFLIPFVIIIFCYTTL-----IHLKSK-----DRWLGYIKAVLLILVIFTI 327

Db 221 FILVWFVW---LFLILLIISYIKGNKLLRISRRSKFPNSGKYATTARNSEFVLFIPTI 277
QY 328 CFAPTNIIIVIHANYVYH-----NTDSLYFMYL-----IALCLGSLNSCLDPPFLYFVM 376
Db 278 CFVP-----YHAFPIYISSQLNVSSCYWKEIVHKTNEIMLVLSFNSCLDPPVNYFLM 330
QY 377 S-----KVVDQL 383
Db 331 SSNIRKIMCOL 341

RESULT 34

P2YR_MOUSE
ID P2YR_MOUSE STANDARD; PRT; 373 AA.
AC P49650;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN P2RY1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Insulinoma;
RX MEDLINE=95298025; PubMed=7779087;
RA Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RL "Cloning of rat and mouse P2Y purinoceptors";
RN Biochem. Biophys. Res. Commun. 211:211-218(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Leon C.;
RT "Thromboresistance in P2Y1 receptor knockout mice";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
CC OF EXTRACELLULAR ATP ON INSULIN SECRETION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; U22829; AAA91302.1; --
DR ENBL; AJ245636; CAB57317.1; --
DR HSDP; P34996; 1DD0.
DR MGI; MGI:105049; P2RY1.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 53 74 1 (POTENTIAL).
FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 127 147 3 (POTENTIAL).
FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 167 188 4 (POTENTIAL).
FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 219 238 5 (POTENTIAL).
FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 266 285 6 (POTENTIAL).
FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 304 328 7 (POTENTIAL).
FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 124 202 BY SIMILARITY.
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42212 MW; 944125E9F4560BB3 CRC64;
Query Match 17.1%; Score 364.5; DB 1; Length 373;
Best Local Similarity 29.5%; Pred. No. 5.7e-14;
Matches 84; Conservative 61; Mismatches 107; Indels 33; Gaps 9;
QY 112 IPAIYILLFVVGVPNSIVTLWKLSLRTK---SISLVIFHTNLAIADLLFCVTLPPKIAIYH 168
Db 54 LPAAVILYFIITGFLGNSVAIMFVFMKPSGISVYMP--NLADFLVILPALIFY 111
QY 169 LGGNNWVGEVNCRIITVYVGYGMYCAILILTCMGINRYLATAHPTVYQKLPKRSFSLM 228
Db 112 FNKTDWIFGDAMCKLQRFIFHVNLYGSLFLTCISAHRYSGVVYPLKSLGRLKKNIAIV 171
QY 229 CGIVVMVFLYMLPFVILKQEVHLVHSEITTCDDVDACESSPFRFYFVSL--AFFGF 286
Db 172 SVLWVLIYVVAISP-ILFYSGTGTRKNKTVTCYDTT---SNDYLSRYFIYSCMTTVAMF 226
QY 287 LIPFVILIFCYTTLTHLKSKD-----RIMLGKAVKALLIIVIFTCFAPTNILV 337
Db 227 CIPLVLILGCGYGLIKVALIYNDLNSPLRRKSIYL-----VIIVLTVFAVSIPFHVMT 281
QY 338 IH-HANYYYHN-----TDSLYFMYLIALCLGSLNSCLDPPFLYFV 375
Db 282 MNLRLRDLDPTEMCDCDNDRVVATYQVTRGLASLNSCVDPILYFL 326

RESULT 35

CLT2_MOUSE
ID CLT2_MOUSE STANDARD; PRT; 309 AA.
AC Q920A1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 2 (CysLTR2).
GN CYSLTR2 OR CYSLTR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=21601669; PubMed=11591709;
RA Hui Y., Yang G., Galczenski H., Figueroa D.J., Austin C.P.,
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Funk C.D.;
RT "The murine cysteinyl leukotriene 2 (CysLTR2) receptor. cDNA and
RT genomic cloning, alternative splicing, and in vitro
RT characterization";
RL J. Biol. Chem. 276:47489-47495(2001).
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system. The rank order of affinities for
CC the leukotrienes is LTC4 = LTD4 >> LTE4.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels, with highest
CC expression in the spleen, thymus and adrenal gland, and lower in
CC the kidney, brain and peripheral blood leukocytes.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cysteinylnyl leukotriene receptor 1 (CysLTR1) (Cysteinylnyl leukotriene D4 receptor) (LTD4 receptor)

DE CysLTR1 OR CysLT1 OR CysLT1R

GN Mus musculus (Mouse)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

OX [1]

RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RP STRAIN=C57BL/6; Tissue=Trachea;

RC PubMed=11226226;

RX Maekawa A., Kanaoka Y., Lam B.K., Austen K.F.;

RA "Identification in mice of two isoforms of the cysteinyl leukotriene 1 receptor that result from alternative splicing.";

RT Proc. Natl. Acad. Sci. U.S.A. 98:2256-2261(2001).

RL [2]

RN SEQUENCE FROM N.A. (ISOFORM 1).

RP STRAIN=C57BL/6; Tissue=Trachea;

RC MEDLINE=21562332; PubMed=11705452;

RX Martin V., Sawyer N., Stocco R., Unett D., Lerner M.R., Abramovitz M., Funk C.D.;

RA "Molecular cloning and functional characterization of murine cysteinyl-leukotriene 1 (CysLT1) receptors.";

RT Biochem. Pharmacol. 62:1193-1200(2001).

RL [3]

RN SEQUENCE FROM N.A. (ISOFORM 1).

RP Tissue=Blood;

RC MEDLINE=21240338; PubMed=11342226;

RX Mollerup J., Jorgensen S.T., Hougard C., Hoffmann E.K.;

RA "Identification of a murine cysteinyl leukotriene receptor by expression in *Xenopus laevis* oocytes.";

RT Biochim. Biophys. Acta 1517:455-459(2001).

RL [4]

RN SEQUENCE FROM N.A. (ISOFORM 2).

RP Tissue=T-cell;

RC Ogasawara H., Izumi T., Shimizu T.;

RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

CC -I- FUNCTION: Receptor for cysteinyl leukotrienes mediating constriction of the microvascular smooth muscle during an inflammatory response. This response is mediated via a G-protein that activates a phosphatidylinositol-calcium second messenger system. The rank order of affinities for the leukotrienes is LTD4 >> LTE4 > LTC4 >> LTB4.

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- ALTERNATIVE PRODUCTS: 2 isoforms; 1/long form (shown here) and 2/short form; are produced by alternative splicing.

CC -I- TISSUE SPECIFICITY: Widely expressed, with higher expression in the lung and skin, intermediate levels in the heart, kidney and stomach and lower levels in several other tissues. Isoform 1 is the most abundant form in all tested tissues.

CC -I- MISCELLANEOUS: MK-571, a selective antagonist, was shown to inhibit eosinophilia, bronchial hyperreactivity and microvascular leakage. Zafirlukast (Accolate) and pranlukast (Onon) were also shown to be selective antagonists.

CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF329272; AA16715.1; -

CC EMBL; AF329272; AA16716.1; -

CC EMBL; AF205830; AA15433.1; -

CC EMBL; AF263370; AA73047.1; -

CC EMBL; AB044087; BAA96809.1; -

CC HSP; P34996; 1DDD.

CC MGD; MGI:1926218; Cysltr1.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.

DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1. Glycoprotein; G-protein coupled receptor; Transmembrane; Glycoprotein; Alternative splicing.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Alternative splicing.

KW DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 44 1 (POTENTIAL).

FT DOMAIN 44 64 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 65 72 2 (POTENTIAL).

FT TRANSMEM 73 93 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 94 121 3 (POTENTIAL).

FT TRANSMEM 122 142 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 143 156 4 (POTENTIAL).

FT TRANSMEM 157 177 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 178 208 5 (POTENTIAL).

FT TRANSMEM 209 229 6 (POTENTIAL).

FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 246 266 7 (POTENTIAL).

FT DOMAIN 267 291 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 292 312 7 (POTENTIAL).

FT DOMAIN 313 352 BY SIMILARITY.

FT DISULFID 111 188 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 15 15 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 19 19 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 184 184 MISSING (IN ISOFORM 2).

FT VARSPLIC 1 13 Y -> D (IN REF. 3).

FT CONFLICT 176 176

FT SEQUENCE 352 AA; 40715 MW; 5BDC94B3F1CD0CAB CRC64;

QY Query Match 16.9%; Score 361; DB 1; Length 352;

Best Local Similarity 30.8%; Pred. No. 8.6e-14;

Matches 103; Conservative 58; Mismatches 129; Indels 44; Gaps 13;

QY 58' GPQNTFEFFPLSDIEGWTGATTIKAECPEDSISTLVHVNNA---TIGYLRSSLSTQVIPA 114

Db 5 GTKQTF---LENNMGTEMLTSL-----INNTCHDTIDEFRN---QVYST 43

QY 115 IVILLFVVGVPSPNIVTLWKLRLT---KSTLVIFHTNLAIADLLFCVTLPPKIAVHLGN 172

Db 44 MYSIVSVGVGFFGNSVLYVL-IKTHKSAFOVMINLAIDLLCVTLPLRVVYVHKG 102

QY 173 NNVFGEVWCRIITVTVFYGNMYCAILILTCMGINRYLATAHPPTYOKLPKRSFLMCGIV 232

Db 103 KWLFGDFLCRLTYALYNLYCSIEFTAMSFRCVAIVFPQINILVTKKARFVIGI 162

QY 233 WVMFLYMLPFVILKQEVHLVHSEITTCDDVVVDACESPSFRFYFVSLAFPGFLIPFVI 292

Db 163 WIFVILTSSPFLMYKS--YQDEKNNTKCFEPQNNQAKYVILHLHVSL-FFGFIIPFVT 219

QY 293 IIFCYTTLHLK--KSKDRILWLYIKAVLLILVI---FTICFAPTNILVTHHANYHYN 347

Db 220 IIVCYTMTLTLKNTKKNMPSRKATGMIVTAAFLVSPMPYHIQRTIH--LHLHS 277

QY 348 ----TDSLVEF---YLIALCLGSLNSCLDPFLYF 374

Db 278 ETRPCDSVLRMQKSVVITLSLAASCCFPDPLLYF 311

RESULT 39

GP34_MOUSE

ID GP34_MOUSE STANDARD; PRT; 375 AA.

AC Q9RIK6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Probable G protein-coupled receptor GPR34.

GN GPR34.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:01:21 ; Search time 17.9779 Seconds
(without alignments)
666.102 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXTXHPVAGSQDIKWIL.....AMARPLXPRRDIWIDHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2130	99.7	407	2	US-08-742-440A-3
2	1433.5	67.6	408	2	US-08-742-440A-6
3	581.5	27.2	395	1	US-08-097-938-5
4	581.5	27.2	395	1	US-08-476-000-5
5	581.5	27.2	395	1	US-08-472-840-5
6	581.5	27.2	395	2	US-08-476-976-5
7	581.5	27.2	395	2	US-08-474-410-5
8	581.5	27.2	395	4	US-08-486-673B-2
9	581.5	27.2	395	4	US-08-486-673B-5
10	581.5	27.2	399	1	US-08-476-000-61
11	581.5	27.2	399	1	US-08-472-840-61
12	581.5	27.2	399	2	US-08-476-976-61
13	581.5	27.2	399	3	US-08-474-410-61
14	581.5	27.2	399	4	US-08-486-673B-61
15	579.5	27.1	395	1	US-08-097-938-2
16	579.5	27.1	395	1	US-08-476-000-2
17	579.5	27.1	395	1	US-08-472-840-2
18	579.5	27.1	395	2	US-08-476-976-2
19	579.5	27.1	395	3	US-08-474-410-2
20	563	26.4	398	1	US-08-097-938-4
21	563	26.4	398	1	US-08-476-000-4
22	563	26.4	398	1	US-08-472-840-4
23	563	26.4	398	2	US-08-476-976-4
24	563	26.4	398	3	US-08-474-410-4
25	563	26.4	398	4	US-08-486-673B-4
26	561.5	26.3	398	4	US-08-486-673B-6
27	560.5	26.2	398	1	US-08-097-938-6

28	560.5	26.2	398	1	US-08-476-000-6	Sequence 6, Appli
29	560.5	26.2	398	1	US-08-472-840-6	Sequence 6, Appli
30	560.5	26.2	398	1	US-08-476-976-6	Sequence 6, Appli
31	560.5	26.2	398	3	US-08-474-410-6	Sequence 6, Appli
32	559	26.2	397	4	US-08-486-673B-63	Sequence 63, Appli
33	553.5	25.9	394	2	US-08-742-440A-8	Sequence 8, Appli
34	551	25.8	397	1	US-08-476-000-63	Sequence 63, Appli
35	551	25.8	397	1	US-08-472-840-63	Sequence 63, Appli
36	551	25.8	397	2	US-08-476-976-63	Sequence 63, Appli
37	551	25.8	397	3	US-08-474-410-63	Sequence 63, Appli
38	509	23.8	425	1	US-07-657-769B-69	Sequence 69, Appli
39	509	23.8	425	1	US-07-789-184-220	Sequence 220, App
40	509	23.8	425	1	US-08-475-263-220	Sequence 220, App
41	509	23.8	425	1	US-08-485-886-220	Sequence 220, App
42	509	23.8	425	2	US-08-477-362-220	Sequence 220, App
43	509	23.8	425	2	US-08-477-134-220	Sequence 220, App
44	509	23.8	425	2	US-08-911-320A-3	Sequence 3, Appli
45	509	23.8	425	2	US-08-742-440A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-742-440A-3
; Sequence 3, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; US-08-742-440A-3
Query Match 99.7%; Score 2130; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 3.3e-155;

Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLTYXQHPVAGSODIKMKILILVAAGLLFLPVTVCOSGINVSDNSAKPTLTITKSFNGGPQ 60
Dy 1 TLTYXQHPVAGSODIKMKILILVAAGLLFLPVTVCOSGINVSDNSAKPTLTITKSFNGGPQ 60
Qy 61 NTPEEPPLSDIEGTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLF 120
Dy 61 NTPEEPPLSDIEGTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLF 120
Qy 121 VVGVPISNIVTLWKLRLTKSISLVIFHTNLAIADLLFCVTLPKFIAYHLNNGNWWFGEVM 180
Dy 121 VVGVPISNIVTLWKLRLTKSISLVIFHTNLAIADLLFCVTLPKFIAYHLNNGNWWFGEVM 180
Qy 181 CRITTVVFGNMYCAIILTCMGINRYLATAHPTTYOKLPKRSFSLMCGIIVWVWFLYM 240
Dy 181 CRITTVVFGNMYCAIILTCMGINRYLATAHPTTYOKLPKRSFSLMCGIIVWVWFLYM 240
Qy 241 LPFVLKQEHVHSEITTCCHVDVDCESPSFRFYFVSLAFPGFLIPFVILFCVTTL 300
Dy 241 LPFVLKQEHVHSEITTCCHVDVDCESPSFRFYFVSLAFPGFLIPFVILFCVTTL 300
Qy 301 IHKLKSDRIWLGVIKAVLLIIVFTICFAPTNIIIVHHANYHHNTDSLYFMYLIALC 360
Dy 301 IHKLKSDRIWLGVIKAVLLIIVFTICFAPTNIIIVHHANYHHNTDSLYFMYLIALC 360
Qy 361 LGSLSCLDPLFYFVMSKVDQLNPXSAMARPLXPRDRDIWEDIHAW 407
Dy 361 LGSLSCLDPLFYFVMSKVDQLNPXSAMARPLXPRDRDIWEDIHAW 407

RESULT 2

US-08-742-440A-6

; Sequence 6, Application US/08742440A

; Patent No. 5892014

; GENERAL INFORMATION:

; APPLICANT: Coughlin, Shaun

; APPLICANT: Ishihari, Hiroaki

; APPLICANT: Connolly, Andrew

; TITLE OF INVENTION: Protease Activated Receptor

; TITLE OF INVENTION: 3 and Uses Thereof

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed, LLP

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/742,440A

; FILING DATE: 30-OCT-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: UCAL/060PAT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400

; TELEFAX: 650 327-3231

; TELEX:

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 408 amino acids

; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
US-08-742-440A-6

Query Match 67.6%; Score 1443.5; DB 2; Length 408;
Best Local Similarity 71.5%; Pred. No. 7.5e-103;
Matches 264; Conservative 45; Mismatches 59; Indels 1; Gaps 1;

Qy 11 GSODIKMKILILVAAGLLFLPVTVCOSGI-NVSDNSAKPTLTITKSFNGGPQNTFEFPPLS 69
Dy 14 GTQVIRKALIFAAGLLLLPTFCOSGMENDTNLAKPTLPKTRGAPPNSFEFPFPS 73
Qy 70 DIEGTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPISNIV 129
Dy 74 ALEGWGTATTVKIKPEESASHLVKNATMGVLTSSLSKLPALYLLVFFVVGVPANAV 133
Qy 130 TLWKLRLTKSISLVIFHTNLAIADLLFCVTLPKFIAYHLNNGNWWFGEVMCRITTVFY 189
Dy 134 TLWMLFFRTSICCTTVFYTNLAIDFLCVTLPPFKIAYHLNNGNWWFGEVLCRATTVIFY 193
Qy 190 GNNYCAIILTCMGINRYLATAHPTTYOKLPKRSFSLMCGIIVWVWFLYMFLPVILKQE 249
Dy 194 GNNYCSILLIACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQE 253
Qy 250 YHLVHSEITTCCHVDVDCESPSFRFYFVSLAFPGFLIPFVILFCVTTLIHKLKSKOR 309
Dy 254 YVLQPDITTCCHVDVNTCESSPFQYFISLAFPGFLIPFVILFCVTTLIHKLKSKOR 313
Qy 310 IWLGYIKAVLLIIVFTICFAPTNIIIVHHANYHHNTDSLYFMYLIALCGLSLNSCLD 369
Dy 314 RNLWYVKASILLIIVFTICFAPSNIILIIHHANYHHNTDGLYFYIYLIALLCLGSLNSCLD 373
Qy 370 PFLYFVMSK 378
Dy 374 PFLYFLMSK 382

RESULT 3

US-08-097-938-5

; Sequence 5, Application US/08097938

; Patent No. 5629174

; GENERAL INFORMATION:

; APPLICANT: SUNDELIN, JOHAN

; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS

; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/097,938

; FILING DATE: 26-JUL-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: MURASHIGE, KATE H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 22803-20006.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 395 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-097-938-5

Query Match      27.2%; Score 581.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPSPNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTITVFLPVWYIIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLPEKTAIYHLNGNNWVGEVNCRIITTVFYGNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHGNNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPRKSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTTCHDVLP-EVLVGD 236
QY 274 RYFYVSLAFFGLIPFVILIFCYTTLI-----HKLKSKDRILWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 292
QY 324 IFTICFAPTNILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLLVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYVYFVSK 347

RESULT 4
US-08-476-000-5
; Sequence 5, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
;
; TOPOLOGY: linear
US-08-476-000-5

Query Match      27.2%; Score 581.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPSPNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTITVFLPVWYIIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLPEKTAIYHLNGNNWVGEVNCRIITTVFYGNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHGNNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPRKSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTTCHDVLP-EVLVGD 236
QY 274 RYFYVSLAFFGLIPFVILIFCYTTLI-----HKLKSKDRILWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 292
QY 324 IFTICFAPTNILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLLVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYVYFVSK 347

RESULT 5
US-08-472-840-5
; Sequence 5, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-840-5
```

[illegible]

RESULT 6
 US-08-476-976-5
 ; Sequence S, Application US/08476976
 ; Patent No. 5874400
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,976
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 395 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-476-976-5

Query Match 27.2%; Score 581.5; DB 2; Length 395;
Best Local Similarity 37.6%; Pred. NO. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY	95	VNNATIGYRRSLSTQVIPAIIYLFPVCVPSPNIWTLKWSLRT-KSISLVIFHTNLATA	153
Dd	S9	IDFESASILTKLTITFLFEVVYIIVFVLGSPGMALWIFLFRTRKKHPAYMYANLALA	118
QY	154	DLLFCVTLPFKIAYHLGNWNWFGEVMCRITTWFFYGNNMYCAILIITCMGINRYLATAHP	213
Dd	119	DLLSVIWFLPKLSYHLHGNWNWYGALCKVLGFPGNMYCSILPMTCLSVQRVWVNP	178
QY	214	FITYQLPKRSFSLMCGIWMVWFLYMLPFVILKOEYHLWHSIEITTHCDVDADACESPSFS	273
Dd	179	MGHPR-KKANIAVGVSIAWLILFLTIPLYVMKOTIYPALNITTCHDVLPE-EVLUGD	236
QY	274	RIFYFUSLAFFGLFPFVIIICYTTLI-----HKLSKDRIWLGYIKAVILLIV	323
Dd	237	MENYFUSLAGVLFPLPALTSAGYUMIKTLRSSAMDEHSEKKRQA----IRLIITVLA	292
QY	324	ITICFAPNTIILVIHHANYHYNTDSLIFYMYLIAI CLGLSNASCDDPFIYFMVK	378
Dd	293	MYFICAPSNLLVWHFYLIKTOROSHVAIYLVVALCITLNACIDPFVYFYVK	347

RESULT 7
US-08-474-410-5
; Sequence 5, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:

CLASSIFICATION: ,
PRIOR APPLICATION DATA: ,
APPLICATION NUMBER: US 08/390,301 ,
FILING DATE: 25-JAN-1995 ,
ATTORNEY/AGENT INFORMATION: ,
NAME: ADLER, REID G. ,
REGISTRATION NUMBER: 30,988 ,
REFERENCE/DOCKET NUMBER: 2803-0006.20 ,
TELECOMMUNICATION INFORMATION: ,
TELEPHONE: (202) 887-1500 ,
TELEFAX: (202) 887-0763 ,
TELEX: 90-4030 ,
INFORMATION FOR SEQ ID NO: 5: ,
SEQUENCE CHARACTERISTICS: ,
LENGTH: 395 amino acids ,
TYPE: amino acid ,
STRANDEDNESS: single ,
TOPOLOGY: linear ,
US-08-474-410-5 ,

```
Query Match          27.2%; Score 581.5; DB 3; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
```

Db 59 IDEFSASILTGLTTFVLPVYIIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118
Qy 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRIITTVVFGNMYCAIILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
Qy 214 FTYQKLPRKSFSLMCMGIVVMVFLYMLPFVILKQEHVHLVHSEITTCDDVVDACESSPSF 273
Db 179 MGHPK-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 236
Qy 274 RFYFVSLAPFGLIPFVILFICVTLLI-----HKLKSKDRILWIGYKAVLLILV 323
Db 237 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 292
Qy 324 IFTICFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYFVYFSK 347

RESULT 8
US-08-486-673B-2
; Sequence 2, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 395
; ORGANISM: Mus musculus
US-08-486-673B-2
Query Match 27.2%; Score 581.5; DB 4; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATICYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTGLTTFVLPVYIIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118
Qy 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRIITTVVFGNMYCAIILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
Qy 214 FTYQKLPRKSFSLMCMGIVVMVFLYMLPFVILKQEHVHLVHSEITTCDDVVDACESSPSF 273
Db 179 MGHPK-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 236
Qy 274 RFYFVSLAPFGLIPFVILFICVTLLI-----HKLKSKDRILWIGYKAVLLILV 323
Db 237 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 292
Qy 324 IFTICFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYFVYFSK 347

RESULT 9
US-08-486-673B-5
; Sequence 5, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan

; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: Protein sequence of C140 receptor
US-08-486-673B-5

Query Match 27.2%; Score 581.5; DB 4; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
Qy 95 VNNATICYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTGLTTFVLPVYIIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118
Qy 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRIITTVVFGNMYCAIILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
Qy 214 FTYQKLPRKSFSLMCMGIVVMVFLYMLPFVILKQEHVHLVHSEITTCDDVVDACESSPSF 273
Db 179 MGHPK-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 236
Qy 274 RFYFVSLAPFGLIPFVILFICVTLLI-----HKLKSKDRILWIGYKAVLLILV 323
Db 237 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 292
Qy 324 IFTICFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYFVYFSK 347

RESULT 10
US-08-476-000-61
; Sequence 61, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995

```
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-000-61

Query Match 27.2%; Score 581.5; DB 1; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGVLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTIVFLPVVYIIIVFVIGLPSNGMALWIFLFRKKHPAVIYMANLALA 122
QY 154 DLLFCVTLPEKTAIYHLNGNNWVGEVWMCRTITVVFYGNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182
QY 214 FTYOKLPKRSFSLMCGIWMVFLVYMLPFVILKOEYHLVHSHITTCDDVVDACESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLVYMKQTIYIPALNITTCDDVLP-EVLVGD 240
QY 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWGLYKAVILLIV 323
Db 241 MENYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 296
QY 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMSK 378
Db 297 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVYSK 351

RESULT 11
US-08-472-840-61
; Sequence 61, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988

; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-61

Query Match 27.2%; Score 581.5; DB 1; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGVLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTIVFLPVVYIIIVFVIGLPSNGMALWIFLFRKKHPAVIYMANLALA 122
QY 154 DLLFCVTLPEKTAIYHLNGNNWVGEVWMCRTITVVFYGNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182
QY 214 FTYOKLPKRSFSLMCGIWMVFLVYMLPFVILKOEYHLVHSHITTCDDVVDACESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLVYMKQTIYIPALNITTCDDVLP-EVLVGD 240
QY 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWGLYKAVILLIV 323
Db 241 MENYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 296
QY 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMSK 378
Db 297 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVYSK 351

RESULT 12
US-08-476-976-61
; Sequence 61, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
```

```
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-61

Query Match      27.2%; Score 581.5; DB 2; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVGVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTTLVFLPVVYIIVFVIGLPSNGMALWIFLRTKKHPPAVIYMANLALA 122

Qy 154 DLLFCVTLTPFKIAYHLNGNNWVGEVWCRITTVVFGNMYCAIILTCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLHNGNNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182

Qy 214 FTYOKLPKRSFSLMCGIVMMVFLYMLPFVILKQEVHLVHSEITTCDDVDVDAESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 240

Qy 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWLGVIKAVLLILV 323
Db 241 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 296

Qy 324 IFTICFAPNTIILVHHANYHHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 297 MYFICFAPSLLLVVHYFLIKTORQSHVYALYLVALCLSTLNSCIDPFPVYFVSK 351

RESULT 13
US-08-474-410-61
; Sequence 61, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-61

Query Match      27.2%; Score 581.5; DB 3; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVGVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTTLVFLPVVYIIVFVIGLPSNGMALWIFLRTKKHPPAVIYMANLALA 122

Qy 154 DLLFCVTLTPFKIAYHLNGNNWVGEVWCRITTVVFGNMYCAIILTCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLHNGNNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182

Qy 214 FTYOKLPKRSFSLMCGIVMMVFLYMLPFVILKQEVHLVHSEITTCDDVDVDAESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 240

Qy 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWLGVIKAVLLILV 323
Db 241 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 296

Qy 324 IFTICFAPNTIILVHHANYHHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 297 MYFICFAPSLLLVVHYFLIKTORQSHVYALYLVALCLSTLNSCIDPFPVYFVSK 351

RESULT 14
US-08-486-673B-61
; Sequence 61, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; PRIOR APPLICATION NUMBER: 1995-06-07
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 61
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-486-673B-61

Query Match      27.2%; Score 581.5; DB 4; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVGVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTTLVFLPVVYIIVFVIGLPSNGMALWIFLRTKKHPPAVIYMANLALA 122

Qy 154 DLLFCVTLTPFKIAYHLNGNNWVGEVWCRITTVVFGNMYCAIILTCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLHNGNNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182

Qy 214 FTYOKLPKRSFSLMCGIVMMVFLYMLPFVILKQEVHLVHSEITTCDDVDVDAESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 240

Qy 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWLGVIKAVLLILV 323
Db 297 MYFICFAPSLLLVVHYFLIKTORQSHVYALYLVALCLSTLNSCIDPFPVYFVSK 351
```

Db 241 MFNYFLSLAIGVFLFPALLTASAYVLMKILTRSSAMDEHSEKKRQRA-----IRLIITVLA 296
Qy 324 IFTICFAPTNILIVHANYHYHNTDSLYFMVLLALCLGSLNSCLDPLFLYFVMSK 378
Db 297 MYFICFAPSNNLLLVVHYFLIKTQROSHVYALYVALCLSLNSCIDPFPVYFVSK 351

RESULT 15

US-08-097-938-2
; Sequence 2, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097.938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-097-938-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;
Qy 95 VNNATIGYLRSSLSSTQVIPAIYILLFVVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTGLTVPVPIVFIIVFVIGLPSNGMALWIFLFTKXKHPAVIYMANLALA 118
Qy 154 DLLFCVTLFPKIAHYHLNGNNVFEVCMCRITTVFYGNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHGNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYVWVNP 178
Qy 214 FTYQKLPRKPSFLMCGIWMVFLYMLPFVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIYAVGSLALWLLIFLVTIPLYVMKTIYIPALNITTTCHDVLP-EVLVGD 236
Qy 274 RPYFVSLAAGFLIPFVIIIFCYTTLIHLKLS-----KDRILWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMKILTRSSAMDEHSEKKRQRA-----IRLIITVLA 292
Qy 324 IFTICFAPTNILIVHANYHYHNTDSLYFMVLLALCLGSLNSCLDPLFLYFVMSK 378
Db 293 MYFICFAPSNNLLLVVHYFLIKTQROSHVYALYVALCLSLNSCIDPFPVYFVSK 347

RESULT 16

US-08-476-000-2
; Sequence 2, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-000-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;
Qy 95 VNNATIGYLRSSLSSTQVIPAIYILLFVVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTGLTVPVPIVFIIVFVIGLPSNGMALWIFLFTKXKHPAVIYMANLALA 118
Qy 154 DLLFCVTLFPKIAHYHLNGNNVFEVCMCRITTVFYGNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHGNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYVWVNP 178
Qy 214 FTYQKLPRKPSFLMCGIWMVFLYMLPFVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIYAVGSLALWLLIFLVTIPLYVMKTIYIPALNITTTCHDVLP-EVLVGD 236
Qy 274 RPYFVSLAAGFLIPFVIIIFCYTTLIHLKLS-----KDRILWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMKILTRSSAMDEHSEKKRQRA-----IRLIITVLA 292
Qy 324 IFTICFAPTNILIVHANYHYHNTDSLYFMVLLALCLGSLNSCLDPLFLYFVMSK 378
Db 293 MYFICFAPSNNLLLVVHYFLIKTQROSHVYALYVALCLSLNSCIDPFPVYFVSK 347

RESULT 17
US-08-472-840-2
; Sequence 2, Application US/08472840


```
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390.301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-0763
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQTQVPAIYILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTGTTLVFLPVVYIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118

Qy 154 DLLFCVTLPPKIAIYHLGNWNVFGEVCMCRITTVVFGNYMCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLGNWNVGEALCKVLIGFFYGNMYCSILFMTCLSVQRVWVIVNP 178

Qy 214 FTYQKLPKRSFSLMCGIWMVWVFLYMLPVLKOEYHLVHSEITTCDDVVDACESPSSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMQTIVIPALNITTCDDVLP-EVLVGD 236

Qy 274 RFYFVSLAPGFLIPFVIIIFCYTTLIHKLK-----KDRWLGYIKAVLLILV 323
Db 237 MFNIFLSAIGVFLPALLTASAYVLMIKTLRSSAMDEHSENKRQRA----IRLIITVLA 292

Qy 324 IFTICFAPNIIIVIHANYHYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVFSK 347

RESULT 18
US-08-476-976-2
; Sequence 2, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
```

```
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-2

Query Match 27.1%; Score 579.5; DB 2; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQTQVPAIYILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTGTTLVFLPVVYIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118

Qy 154 DLLFCVTLPPKIAIYHLGNWNVFGEVCMCRITTVVFGNYMCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLGNWNVGEALCKVLIGFFYGNMYCSILFMTCLSVQRVWVIVNP 178

Qy 214 FTYQKLPKRSFSLMCGIWMVWVFLYMLPVLKOEYHLVHSEITTCDDVVDACESPSSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMQTIVIPALNITTCDDVLP-EVLVGD 236

Qy 274 RFYFVSLAPGFLIPFVIIIFCYTTLIHKLK-----KDRWLGYIKAVLLILV 323
Db 237 MFNIFLSAIGVFLPALLTASAYVLMIKTLRSSAMDEHSENKRQRA----IRLIITVLA 292

Qy 324 IFTICFAPNIIIVIHANYHYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVFSK 347

RESULT 19
US-08-474-410-2
; Sequence 2, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
```

NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-410-2

Query Match 27.1%; Score 579.5; DB 3; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;
QY 95 VNNATIGYRLSSLTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
DB 59 IDEFSASILTGLTTFVPLVYIIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLFPKIAVHLNGNWMVGEVCMCRITTVFYGNMYCAIILLTCMGINRYLATAHP 213
DB 119 DLLSVINFPFLKISYVHLHNGNWMVGEALCKVLIGFYGNMYCSILLFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPKRSFLMCMGIVWVMFLYMLPFVILKQEHVHSEITTCDDVVDACESPSF 273
DB 179 MCHPR-KKANTAVGSLALWLLIFLVTIPLYVMQTIVIPALNITTCHDVLPE-EVLVGD 236
QY 274 RPYFVSLAAGFLIPFVIIIFCYTTLIHLKS-----KDRIWLGVIKAVLILV 323
DB 237 MFNYFLSLAIGVFLPALLTASAYVLMKTLRSSAMDESHSNKQRA----IRLIITVLA 292
QY 324 IFTICFAPTNILVTHANYHHNTDSLYFMYLALCLGSLNSCLDPLFYFVMSK 378
DB 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYVALCLSLNSCLDPLFYFVSK 347

RESULT 20
US-08-97-938-4
Sequence 4, Application US/08097938
Patent No. 5629174
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,938
FILING DATE: 26-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22803-20006.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-097-938-4

Query Match 26.4%; Score 563; DB 1; Length 398;
Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;
QY 49 TLTIKSPNGQPNTFEFPFLSDIEGW---TGATTIKACPEDSISTLHVNATIGYRLS 105
DB 21 TLVFLSCTGTNRSSKGRSLIGKVDGTSHTVGKVTVE-----TVFSDEFSSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIAADLLFCVTLFPK 164
DB 73 KLITVFLPIVYITVIVVUGLPSNGMALWVFLFRKKKHPAVIYMANLALADLLSVIFPLK 132
QY 165 IAYHLNGNWMVGEVCMCRITTVFYGNMYCAIILLTCMGINRYLATAHPPTYOKLPKRSF 224
DB 133 IAVHHGNWNYGEALCNVLIGFYGNMYCSILLFMTCLSVQRYWVIVNPMGHSR-KKANI 191
QY 225 SLUMCGIVWVMFLYMLPFVILKQEHVHSEITTCDDVVDACESPSF-----RFYFVS 280
DB 192 AIGISLAIMLLILLVPIYVVKOTIFIPALNITTCHDVL-----PEQLLVGDMFYFLS 246
QY 281 LAFFGFLIPFVIIIFCYTTLIHLKS-----KDRIWLGVIKAVLILVIFTICFAPTNI 334
DB 247 LAIGVFLPFAFLTASAYVLMIRLRSAMDENSEKKRKAIKLIVTVLAMYLCIFTPSNL 306
QY 335 ILVHHANYHHNTDSLYFMYLALCLGSLNSCLDPLFYFVMS 377
DB 307 LLVVHVFLLKSQCSQSHVYALYVALCLSLNSCLDPLFYFVS 349

RESULT 21
US-08-476-000-4
Sequence 4, Application US/08476000
Patent No. 5716789
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-000-4

Query Match 26.4%; Score 563; DB 1; Length 398;

Best Local Similarity 33.5%; Pred. No. 1e-35;

Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

```
QY 49 TLTIKSNQGPQNTFEFPFLSDIEGW---TGATTTKAECPEDSISTLHVNATIGVLR 105
Db 21 TLVFLSCTGNRSKGRSLGKVDGTSHTVKGVTVE-----TVFSVDEFSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KSILVIFHTNLAIADLLFCVTLPPK 164
Db 73 KLTVFLPIVYIVFVVGLPNSGMAVFLFRTKKHPAVIYMANLADLLSVIWFPLK 132
QY 165 IAYHLGNMNVGEVCMCRITTVVYGNMYCAILILTCMGINRYLATAHPTTYOKLPKRSF 224
Db 133 IAYHIGNNMIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYVWVIVNPMGHSR-KKANI 191
QY 225 SLLMCGIIVWVFLYMLPFVILKQYHLVHSEITTCDDVDVDAESPSSF-----RFYFVVS 280
Db 192 AIGISLAIWLILVITPLYVVKQTFIPALNITTCDDVL-----PEQLLVGDMFNFLS 246
QY 281 LAFFGFLIPFVIIIFCYTTLIHKLKS-----KDRWLGVKAVLLILVIFTICFAPTNI 334
Db 247 LAIGVFLFPAFLTASAYVLMIRMSAMDENSEKRRKRAIKLIVTVLAMYLICFTPSNL 306
QY 335 ILVIHANYYYHNTDSLYFMYLIACLGSLNSCLDPFLYFVMS 377
Db 307 LLVWHYFLIKSQGQSHVYALYVALCLSTLNSCIDPEVYFVVS 349
```

RESULT 22

US-08-472-840-4

Sequence 4, Application US/08472840

Patent No. 5763575

GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN

APPLICANT: SCARBOROUGH, ROBERT M.

TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,840
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-4

Query Match 26.4%; Score 563; DB 1; Length 398;

Best Local Similarity 33.5%; Pred. No. 1e-35;

Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

```
QY 49 TLTIKSNQGPQNTFEFPFLSDIEGW---TGATTTKAECPEDSISTLHVNATIGVLR 105
Db 21 TLVFLSCTGNRSKGRSLGKVDGTSHTVKGVTVE-----TVFSVDEFSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KSILVIFHTNLAIADLLFCVTLPPK 164
Db 73 KLTVFLPIVYIVFVVGLPNSGMAVFLFRTKKHPAVIYMANLADLLSVIWFPLK 132
QY 165 IAYHLGNMNVGEVCMCRITTVVYGNMYCAILILTCMGINRYLATAHPTTYOKLPKRSF 224
Db 133 IAYHIGNNMIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYVWVIVNPMGHSR-KKANI 191
QY 225 SLLMCGIIVWVFLYMLPFVILKQYHLVHSEITTCDDVDVDAESPSSF-----RFYFVVS 280
Db 192 AIGISLAIWLILVITPLYVVKQTFIPALNITTCDDVL-----PEQLLVGDMFNFLS 246
QY 281 LAFFGFLIPFVIIIFCYTTLIHKLKS-----KDRWLGVKAVLLILVIFTICFAPTNI 334
Db 247 LAIGVFLFPAFLTASAYVLMIRMSAMDENSEKRRKRAIKLIVTVLAMYLICFTPSNL 306
QY 335 ILVIHANYYYHNTDSLYFMYLIACLGSLNSCLDPFLYFVMS 377
Db 307 LLVWHYFLIKSQGQSHVYALYVALCLSTLNSCIDPEVYFVVS 349
```

RESULT 23

US-08-476-976-4

Sequence 4, Application US/08476976

Patent No. 5874400

GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN

APPLICANT: SCARBOROUGH, ROBERT M.

TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500


```
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-486-673B-4

Query Match      26.4%; Score 563; DB 4; Length 398;
Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

Qy 49 TLTIKSNFGGPNTEEFPSLSDIEGW---TGATTIKAECPEDSISTLHVNNATIGVLR 105
Db 21 TLVFLSTGNTNRSSKGRSLGKVDGTSHTVKGVTVE-----TVFSVDEFSASVLTG 72

Qy 106 SLSTQVIPAIVILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIADLLFCVTLPPK 164
Db 73 KLTTFELPIVTVIVVVGPSNGMALWFLFRTKKHPAVIYMANLADLLSVIWFPLK 132

Qy 165 IAYHLGNVNVFGEVMCRTITVVFYGNMYCAILITCMGINRYLATAPHTYQKLPKRSF 224
Db 133 IAYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSR-KKANI 191

Qy 225 SLLMCGIVVMVFLYMLPFVILKQYHLVHSEITTCCHVDVDCSPSSF----RFYFVS 280
Db 192 AIGSLALWLLVLTPIYVVKOTIFIPALNITTCCHDVL-----PQLLVGDMFNFLS 246

Qy 281 LAFFGLPIPFVILIFCYTTLIHKLK-----KDRILWGLYIKAVLLILVIFTICFAPTNI 334
Db 247 LAIGVFLFPALFTASAYVLMRLRSSAMDENSEKKRAIKLIVTVLWMLYLCFTSNL 306

Qy 335 ILVTHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 307 LLVVHYFLIKSQGSHVYALVALCLSLTNSCIDPFVYFVS 349

RESULT 26
US-08-486-673B-6
; Sequence 6, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486.673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Protein sequence from cDNA
US-08-486-673B-6

Query Match      26.3%; Score 561.5; DB 4; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.4e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGVLRSSLTQVIPAIVILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAI 153
Db 62 VDEFSASVLTGKLTIVFLPIVYIIVFVVGPSNGMALWFLFRTKKHPAVIYMANLALA 121

Qy 154 DLLFCVTLPPFKIAYHLGNVNVFGEVMCRTITVVFYGNMYCAILITCMGINRYLATAP 213
Db 122 DLLSVIWFPLKIAHYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 181

Qy 214 FTYQKLPKRSFSLMCGIVVMVFLYMLPFVILKQYHLVHSEITTCCHVDVDCSPSSF 273
Db 182 MHSR-KKANIAIGISLAIWLLVLTPIYVVKOTIFIPALNITTCCHDVL-----PEQL 235

Qy 274 ---REVYFVSFLAFFGLPIPFVILIFCYTTLIHKLK-----KDRILWGLYIKAVLLILV 323
Db 236 LVGDMFNFLSALIGVFLFPALFTASAYVLMRLRSSAMDENSEKKRAIKLIVTVLA 295

Qy 324 IFTICFAPTNIILVTHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 296 MYLICFIPSNLLVHVHYFLIKSQGSHVYALVALCLSLTNSCIDPFVYFVS 349

RESULT 27
US-08-097-938-6
; Sequence 6, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-097-938-6

Query Match      26.2%; Score 560.5; DB 1; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.6e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGVLRSSLTQVIPAIVILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAI 153
Db 62 VDEFSASVLTGKLTIVFLPIVYIIVFVVGPSNGMALWFLFRTKKHPAVIYMANLALA 121

Qy 154 DLLFCVTLPPFKIAYHLGNVNVFGEVMCRTITVVFYGNMYCAILITCMGINRYLATAP 213
Db 122 DLLSVIWFPLKIAHYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 181

Qy 214 FTYQKLPKRSFSLMCGIVVMVFLYMLPFVILKQYHLVHSEITTCCHVDVDCSPSSF 273
Db 182 MHSR-KKANIAIGISLAIWLLVLTPIYVVKOTIFIPALNITTCCHDVL-----PEQL 235
```

Db 182 MHSR-KKANAIGISLAIWLLILLVTPLYVVKQTIFIPALNITTTCHDVL-----PEQV 235
Qy 274 ----RFFYVSLAFGLIPVPIIFCVTTLIHLKLS-----KDRWLGVYKAVLLILV 323
Db 236 LVGDMFNYSLSAIGVFLFPFAFLTASAYVLMIRLSSAMENSEKKRKAUKLIVTLA 295
Qy 324 IFTICFAPTNILVIHHANYHYNTDSLYFMYLIALCLGSLNSCLDPLFYFVMS 377
Db 296 MYLICFIPSNLLLVVHFLIKSQSQSHVYALYIIVALCLSTLNSCIDPFVYFVS 349

RESULT 28
US-08-476-000-6
; Sequence 6, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-476-000-6

Query Match 26.2%; Score 560.5; DB 1; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.6e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSTRT-KSISLVIFHTNLAIA 153
Db 62 VDEFSASVLTGKLTITVFLPIVYIIVFVGLPSNGMALWVFLFRTKKHPAIVYMANLALA 121
Qy 154 DLLFCVTLTPFKIAYHLNGNNVFEVCMRITTVVFGYGMVCAIILTCMGINRYLATAHP 213
Db 122 DLLSVWFPFKIAYHIGNNWIYGEALCNVLGFFYGMVCSILFMTCLSVQRYWVIVNP 181
Qy 214 FTYOKLPKRSFLMCGIWMVWVFLYMLPFVILKQEHVHSEITTTCHDVVDACESPSS 273
Db 122 DLLSVWFPFKIAYHIGNNWIYGEALCNVLGFFYGMVCSILFMTCLSVQRYWVIVNP 181
Qy 214 FTYOKLPKRSFLMCGIWMVWVFLYMLPFVILKQEHVHSEITTTCHDVVDACESPSS 273
Db 182 MHSR-KKANAIGISLAIWLLILLVTPLYVVKQTIFIPALNITTTCHDVL-----PEQV 235
Qy 274 ----RFFYVSLAFGLIPVPIIFCVTTLIHLKLS-----KDRWLGVYKAVLLILV 323

Db 236 LVGDMFNYSLSAIGVFLFPFAFLTASAYVLMIRLSSAMENSEKKRKAUKLIVTLA 295
Qy 324 IFTICFAPTNILVIHHANYHYNTDSLYFMYLIALCLGSLNSCLDPLFYFVMS 377
Db 296 MYLICFIPSNLLLVVHFLIKSQSQSHVYALYIIVALCLSTLNSCIDPFVYFVS 349

RESULT 29
US-08-472-840-6
; Sequence 6, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-840-6

Query Match 26.2%; Score 560.5; DB 1; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.6e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSTRT-KSISLVIFHTNLAIA 153
Db 62 VDEFSASVLTGKLTITVFLPIVYIIVFVGLPSNGMALWVFLFRTKKHPAIVYMANLALA 121
Qy 154 DLLFCVTLTPFKIAYHLNGNNVFEVCMRITTVVFGYGMVCAIILTCMGINRYLATAHP 213
Db 122 DLLSVWFPFKIAYHIGNNWIYGEALCNVLGFFYGMVCSILFMTCLSVQRYWVIVNP 181
Qy 214 FTYOKLPKRSFLMCGIWMVWVFLYMLPFVILKQEHVHSEITTTCHDVVDACESPSS 273
Db 182 MHSR-KKANAIGISLAIWLLILLVTPLYVVKQTIFIPALNITTTCHDVL-----PEQV 235
Qy 274 ----RFFYVSLAFGLIPVPIIFCVTTLIHLKLS-----KDRWLGVYKAVLLILV 323
Db 236 LVGDMFNYSLSAIGVFLFPFAFLTASAYVLMIRLSSAMENSEKKRKAUKLIVTLA 295

Qy	324	I	F	T	C	P	A	P	T	N	I	I	V	H	A	N	Y	Y	H	N	T	D	S	L	Y	F	M	Y	L	I	A	L	C	G	S	N	S	C	L	D	P	F	L	Y	F	V	M	S	377
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	296	M	Y	L	I	C	I	P	S	N	L	L	V	V	H	F	L	K	S	Q	G	S	H	V	A	L	Y	I	V	A	L	C	L	S	T	N	S	C	I	D	P	F	Y	F	V	S	349		

RESULT 30
 US-08-476-976-6
 ; Sequence 6, Application US/08476976
 ; Patent No. 5874400
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,976
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 904030
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-476-976-6

RESULT 31
 US-08-474-410-6
 ; Sequence 6, Application US/08474410
 ; Patent No. 6043212
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,410
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-474-410-6

RESULT 32

US-08-486-673B-63
; Sequence 63, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; PRIOR FILING DATE: 1995-06-07
; PRIOR FILING DATE: 1993-07-26
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-486-673B-63

Query Match 26.2%; Score 559; DB 4; Length 397;
Best Local Similarity 32.4%; Pred. No. 2.1e-35;
Matches 119; Conservative 72; Mismatches 138; Indels 38; Gaps 8;
Qy 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGSPQNTFEFPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPSTNITVLKLSLRT-KS 140
Db 56 E-----TVFSVDFESASVLGKLTIVFLPIVTVTFVAVGLPSNGMALWVFLFRYKK 107
Qy 141 ISLVIFHTLAIADLLFCVTLPPFKIAYHLNGNNWVFGEVNCRIITVVYFNGMYCAILIT 200
Db 108 HPVIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYCSILFMT 167
Qy 201 CMGINRYLATAPTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEHVHSEITTC 260
Db 168 CLSVQRYWIVNPMGHSR-KKANAIGISLAIWLLTLVTVIPLVYVVKQTIFIPALNITTC 226
Qy 261 HDVVDAECPSSP---RFYFVSFAFFGLPIPVIIIFCVTTLIHLKS-----KDRI 310
Db 227 HDVL-----PQLLVGDMFNFLSLAIGVFLFPALFASAVLMIRLRSAMDENSEKK 281
Qy 311 WLGYIKAVLLILVITTCFAPTNIIIVHANYHHNTSDLSYFMYIALCIGSLNSCLDP 370
Db 282 RKRAIKLIVTVLGMVLCFTPSNLLLVVHYFLIKSQGSHVYALYIVALCLSTLNSCIDP 341
Qy 371 FLYFVMS 377
Db 342 VYFVFS 348

RESULT 33

US-08-742-440A-8
; Sequence 8, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,440A
FILING DATE: 30-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: UCAL/060PAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-440A-8

Query Match 25.9%; Score 553.5; DB 2; Length 394;
Best Local Similarity 32.2%; Pred. No. 5.5e-35;
Matches 118; Conservative 73; Mismatches 136; Indels 39; Gaps 8;
Qy 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGSPQNTFEFPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPSTNITVLKLSLRT-KS 140
Db 56 E-----TVFSVDFESASVLGKLTIVFLPIVTVTFVAVGLPSNGMALWVFLFRYKK 107
Qy 141 ISLVIFHTLAIADLLFCVTLPPFKIAYHLNGNNWVFGEVNCRIITVVYFNGMYCAILIT 200
Db 108 HPVIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYCSILFMT 167
Qy 201 CMGINRYLATAPTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEHVHSEITTC 260
Db 168 CLSVQRYWIVNPMGHSR-KKANAIGISLAIWLLTLVTVIPLVYVVKQTIFIPALNITTC 226
Qy 261 HDVVDAECPSSP---RFYFVSFAFFGLPIPVIIIFCVTTLIHLKS-----KDRI 311
Db 227 HDVL-----PQLLVGDMFNFLSLAIGVFLFPALFASAVLMIRLRSAMDENSEKK 279
Qy 312 LGYIKAVLLILVITTCFAPTNIIIVHANYHHNTSDLSYFMYIALCIGSLNSCLDP 371
Db 280 KRAIKLIVTVLGMVLCFTPSNLLLVVHYFLIKSQGSHVYALYIVALCLSTLNSCIDP 339
Qy 372 LYFVMS 377
Db 340 VYFVFS 345

RESULT 34

US-08-476-000-63
; Sequence 63, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5550
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/98/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-000-63

Query Match	25.8%;	Score 551;	DB 1;	Length 397;
Best Local Similarity	32.2%;	Pred. No. 8.5e-35;		
Matches 118;	Conservative	72;	Mismatches 139;	Indels 38; Gaps 8;
QY	22	LVAAGLLFLPVTCQSGINVSNDNSAKPTLTKSFNGGPGQNTPEEPPLSDIEGWTGATTI	81	
DB	9	LLGRAILLAAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TKGVTV	55	
QY	82	KAECPEDSISTLHYNNATIQYLRSSTQVIPAIIYLLFVVGVPSPNIVTLWKLSLRT-KS	140	
DB	56	E-----TVFSVDEFSASVLAGKLTTFVPIVTVIVFAVGLPSNGMALVFLPRTKKK	107	
QY	141	ISLVFHTNLAIADLLFCVTLPKPIAYHLNGNNWVGEVWCRITTVVFGNMYCAILLT	200	
DB	108	HPAVIYMANLADLLSVIWFPLKIAYIHGNNWIIYGEALCNVLGFFFYRNNMYCSILFMT	167	
QY	201	CMGINRYLATAPPTYOKLPRKSPFLMCGIVVMVFLYMLPFVILKQBYHLVHSEITTC	260	
DB	168	CLSVQRYWVIVNPNHGSR-KKANIAIGISLAIWLLTLLVTIPLVYVKQTFIPALNIITC	226	
QY	261	HDVVDACESPSF-----RFYFVSVLAFPGFLIPFVIIICFYTTLIHUKS-----KDRI	310	
DB	227	HDVL-----PQLLVGDMEFYFLSLAIGVLFPAFLTASAYVIMRLMRLESSAMDNSEKK	281	
QY	311	WLGYIKAVLLTILVFTICFAPTNILIVIHANYYYHNTDSLAFMYLIALCLGLSNCLDP	370	
DB	282	RKRAIKLIVTGLGMYLICFTPSNLLLVVHYFLTKSQGQSHVALYIVACLSTLNSCIDP	341	
QY	371	FLYFVMS	377	
DB	342	FVYFEVS	348	

RESULT 35
US-08-472-840-63
; Sequence 63, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN

RESULT 36

US-08-476-976-63
; Sequence 63, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-63

Query Match 25.8%; Score 551; DB 2; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.5e-35;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;

Qy 22 LVAAGLLFLPVTVCQSGINVSNDNAKPTLTIKSFNGGPQNTFEFFPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KS 140
Db 56 E-----TVFSVDEFSASVLAKLTTFVLPVITVFAVGLPSNGMALWVFLFRKK 107
Qy 141 ISLVIFHTNLAIADLLFCVTLPEKIAVHLNNGNWFGEVWCRTTTFVFGNMYCAIILT 200
Db 108 HPVIVMANLADLLSVIWFPLKIAVHNGNWIYGEALCNVLIGFFYRNMYCSILFMT 167
Qy 261 HDVVDAESPSSP-----RFYFVSLAFPGFLIPFVIIIFCYTTLIHLKLS-----KDRI 310
Db 227 HDVL-----PEQLLVGDMFNFLSLAIGVFLPFAFLTASAYVLMIRLSSAMDNSEKK 281
Qy 311 WLGVIKAVLLIIVTICFAPTNILVIHANYHYHTDSLYFMYLALCLGSLNSCLDP 370
Db 282 RKRAIKLIVTGLMGLICFTSNLLVHVHFLIKSQSQSHVYALYIIVALCLTSLNSCIDP 341
Qy 371 FLYFVMS 377

Db 342 FVYFVS 348
RESULT 37
US-08-474-410-63
; Sequence 63, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-63

Query Match 25.8%; Score 551; DB 3; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.5e-35;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;

Qy 22 LVAAGLLFLPVTVCQSGINVSNDNAKPTLTIKSFNGGPQNTFEFFPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KS 140
Db 56 E-----TVFSVDEFSASVLAKLTTFVLPVITVFAVGLPSNGMALWVFLFRKK 107
Qy 141 ISLVIFHTNLAIADLLFCVTLPEKIAVHLNNGNWFGEVWCRTTTFVFGNMYCAIILT 200
Db 108 HPVIVMANLADLLSVIWFPLKIAVHNGNWIYGEALCNVLIGFFYRNMYCSILFMT 167
Qy 201 CMGINRYLATAHPTTYOKLPKRSFSLMCGIWMVWVFLYMLPEVILKQEHVHSEITTC 260
Db 168 CLSVQRWVIVNPMGHSR-KKANIAIGISLAIWLLTLLVTIPLVYVVKQTIFIPALNITTC 226
Qy 261 HDVVDAESPSSP-----RFYFVSLAFPGFLIPFVIIIFCYTTLIHLKLS-----KDRI 310
Db 227 HDVL-----PEQLLVGDMFNFLSLAIGVFLPFAFLTASAYVLMIRLSSAMDNSEKK 281

QY 311 WLGVKAVLLVLTFTICFAPTNIILVHHANYHNTDLSLYFWYIALCLGSLNSCLDP 370
 Db 282 RKRAKLIVLVGLYICFTPSNLLLVVHYFLIKSQGSHVYALYIIVALCLSTINSICDP 341
 QY 371 FLYFVMS 377
 Db 342 FVYFVS 348

RESULT 38

US-07-657-769B-69
 ; Sequence 69, Application US/07657769B
 ; Patent No. 5256766
 ; GENERAL INFORMATION:
 ; APPLICANT: COUGHLIN, SHAUN R.
 ; TITLE OF INVENTION: RECOMBINANT RECEPTOR AND RELATED
 ; PHARMACEUTICALS
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: IRELL & MANELLA
 ; STREET: 545 MIDDLEFIELD ROAD, SUITE 200
 ; CITY: MENLO PARK
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/657,769B
 ; FILING DATE: 19910219
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 2000-0502.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-327-7250
 ; TELEFAX: 415-327-2951
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 69:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 425 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-657-769B-69

Query Match 23.8%; Score 509; DB 1; Length 425;
 Best Local Similarity 31.2%; Pred. No. 1.5e-31;
 Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
 QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTFEEFPLSDI 71
 Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRNPNDKYEPFWEDEE 60
 QY 72 EGMTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYILLFVVGVPNS 127
 Db 61 KNESGLTEYRLVSVINKSSPLQKLPAPFISEDAGYLTSSWLTLPVPSVYTGVFVVSPLN 120
 QY 128 I--VTWLKLSLRTKYSISLVIFHTNLAIADLLFCVTLFPFKIAYHLNGNWNVFGVMCRITT 185
 Db 121 IMAIVVFLKMKYKK-PAVVYMLHLATADVLFSVLPFKISYFSGSDWQFGSELCHFTV 179
 QY 186 VVFGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVVMFLYMLP 242
 Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVWP 236
 QY 243 FVLKQEHVHSEITTCCHVDVDAVSPSFRFYFVSLAFAFGFLIPFVIIIFCYTTLIH 302
 Db 237 LVLKEQTIQVPLNITTCCHVDVDAVSPSFRFYFVSLAFAFGFLIPFVIIIFCYTTLIH 295

QY 303 KLKS-----KDRIMLGVIKAVLL---ILVIFTICFAPTNIILVHHANYHNTDLSLYFWYIALCLGSLNSCLDP 370
 Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIFCGTNNVLLIAHY-SFLSHTSTTE 347
 QY 350 SLYFWYIALCLGSLNSCLDPFLYFVMS 377
 Db 348 AAYFAYLLCVCVSSISSICDPLIYYAS 375

RESULT 39

US-07-789-184-220
 ; Sequence 220, Application US/07789184
 ; Patent No. 5688768
 ; GENERAL INFORMATION:
 ; APPLICANT: COUGHLIN, SHAUN R.
 ; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
 ; RELATED PHARMACEUTICALS
 ; NUMBER OF SEQUENCES: 223
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/789,184
 ; FILING DATE: 19911107
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22000-20502.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 34-0154
 ; INFORMATION FOR SEQ ID NO: 220:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 425 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-789-184-220

Query Match 23.8%; Score 509; DB 1; Length 425;
 Best Local Similarity 31.2%; Pred. No. 1.5e-31;
 Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
 QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTFEEFPLSDI 71
 Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRNPNDKYEPFWEDEE 60
 QY 72 EGMTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYILLFVVGVPNS 127
 Db 61 KNESGLTEYRLVSVINKSSPLQKLPAPFISEDAGYLTSSWLTLPVPSVYTGVFVVSPLN 120
 QY 128 I--VTWLKLSLRTKYSISLVIFHTNLAIADLLFCVTLFPFKIAYHLNGNWNVFGVMCRITT 185
 Db 121 IMAIVVFLKMKYKK-PAVVYMLHLATADVLFSVLPFKISYFSGSDWQFGSELCHFTV 179
 QY 186 VVFGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVVMFLYMLP 242
 Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVWP 236
 QY 243 FVLKQEHVHSEITTCCHVDVDAVSPSFRFYFVSLAFAFGFLIPFVIIIFCYTTLIH 302

Db 237 LVLKEQTIQVGLNITTCDDVNETLLEGGYAYVFSAVF-FFVPLIISTVCYVSIIR 295
QY 303 KLKS-----KDRWLGVYKAVLL---ILVFTICFAPTNIILVHHANYHH-NTD 349
Db 296 CLSSAVANRKKSR-----ALFLSAAVFCIFICGPTNVLLIAHY-SPLSHTSTTE 347
QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 40

US-08-475-263-220
; Sequence 220, Application US/08475263
; Patent No. 5759994

; GENERAL INFORMATION:

; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/475,263

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: MURASHIGE, KATE H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 22000-20502.03

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 220:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 425 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-475-263-220

Query Match 23.8%; Score 509; DB 1; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.5e-31;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCOSGINVDSNAKP-----TLTIKSP-NGGPQNTFFEEPLSDI 71
Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDPKSLRNPNPKYEPFWEDEE 60
QY 72 EGWTGAT---TTIKACPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPSPN 127
Db 61 KNESGLTEYRLVLSINKSPLOKQLPAFISEDAGSLTSSWLTLPVPSVYTGVPVLSPLN 120
QY 128 I--VTLNKLRLTKSISLVIFPHNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVMCRITT 185
Db 121 IMAIVFILKMKVKK-PAVVVYMLHLATADLVFVSVLPFKISYYPSGSDWQFGSELCRFVT 179
QY 186 VVFYGNMYCAILLTCHGINRYLATAHP---FTYQKLPKRSFSLMCGIIVWVFLYMLP 242
Db 180 AAFYCNMYAIIIMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236

QY 243 FVILKOEYHLVHSEITTCDDVWDACESPSSRFYFVSLAFFGFLIPFVILIFCYTTLIH 302
Db 237 LVLKEQTIQVGLNITTCDDVNETLLEGGYAYVFSAVF-FFVPLIISTVCYVSIIR 295
QY 303 KLKS-----KDRWLGVYKAVLL---ILVFTICFAPTNIILVHHANYHH-NTD 349
Db 296 CLSSAVANRKKSR-----ALFLSAAVFCIFICGPTNVLLIAHY-SPLSHTSTTE 347
QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

Search completed: June 24, 2003, 12:07:18

Job time : 20.9779 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:52:41 ; Search time 51.9362 Seconds
(without alignments)
1044.224 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXQHPVAGSODIKMKIL.....AMARPLXPRRDIWIHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1940	90.8	369	19 AAW51405	Mouse protease-act
2	1425.5	66.7	374	19 AAW51406	Human protease-act
3	581.5	27.2	399	16 AAR66922	Murine Cl40 recept
4	581.5	27.2	399	17 AAW01954	Murine Cl40 recept
5	574.5	26.9	395	16 AAR66920	Murine Cl40 recept
6	572.5	26.8	395	17 AAW01952	Murine Cl40 recept
7	569	26.6	397	21 AAB35641	Human PAR-2 protei
8	563	26.4	398	16 AAR66921	Human Cl40 recepto
9	563	26.4	398	17 AAW01953	Human Cl40 recepto
10	553.5	25.9	394	19 AAW51408	Human protease-act

11	551	25.8	397	16 AAR66923	Human Cl40 recept
12	551	25.8	397	17 AAW01955	Human Cl40 recepto
13	509.5	23.9	385	20 AAY50139	Mutant human prote
14	509.5	23.9	408	20 AAY50137	FLAG epitope-tagge
15	509	23.8	425	13 AAR27240	Human thrombin rec
16	509	23.8	425	19 AAW51407	Human protease-act
17	509	23.8	425	20 AAY49570	Human thrombin rec
18	509	23.8	425	23 AAG80697	Human thrombin pro
19	509	23.8	425	23 AAE17032	Human thrombin pro
20	509	23.8	426	21 AAY45035	Human thrombin rec
21	507.5	23.8	385	20 AAY15082	Human protease-act
22	507.5	23.8	385	20 AAY50135	Human protease-act
23	507.5	23.8	385	21 AAY45036	Human protease act
24	507.5	23.8	385	22 AAB47623	Human PAR4. Homo
25	507	23.7	425	15 AAR60698	Fragment of the hu
26	506.5	23.7	385	20 AAY50138	Mutant human prote
27	502	23.5	892	18 AAW16314	Human thrombin rec
28	501	23.5	425	23 AEG35300	Human PAR1 type th
29	494	23.1	359	21 AAY71304	Human orphan G pro
30	494	23.1	359	21 AAB02838	Human G protein co
31	493	23.1	396	20 AAY15081	Mouse protease-act
32	493	23.1	359	21 AAY84815	A human G-protein
33	493	23.1	359	22 AAG78530	G-protein coupled
34	493	23.1	359	22 AAG80966	Human nPCR5 #2.
35	493	23.1	359	22 AAB62285	Human G-protein co
36	489	22.9	359	21 AAY69485	Amino acid sequenc
37	485.5	22.7	402	23 AAG35298	Human PAR1 type th
38	484.5	22.7	371	23 AAG35299	Human PAR1 type th
39	444.5	20.8	359	19 AAW69598	Mouse G-protein co
40	431.5	20.2	361	15 AAR54080	Epstein Barr virus
41	431.5	20.2	361	19 AAW53623	Epstein Barr virus
42	431.5	20.2	361	21 AAY90630	Human G protein-co
43	429	20.1	374	22 AAE04390	Turkey P2Y nucleot
44	428.5	20.1	361	21 AAY90664	Human mutant G pro
45	407	19.1	322	15 AAR48715	G-protein coupled

ALIGNMENTS

RESULT 1
AAW51405
ID AAW51405 standard; Protein; 369 AA.
XX
AAW51405;
XX
DT 12-OCT-1998 (first entry)
XX
Mouse protease-activated receptor 3 (PAR3).
XX
DE DE
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis.
XX
Mus sp.
XX
WO9818456-A1.
XX
PD 07-MAY-1998.
XX
PF 29-OCT-1997; 97WO-US19732.
XX
PR 30-OCT-1996; 96US-0742440.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Connolly A, Coughlin SR, Ishihara H;
XX
DR WPI; 1998-271905/24.
XX
N-PSDB; AAV07372.

PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.

PS Claim 3; Page 39-40; 74pp; English.

XX This polypeptide comprises mouse protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372). The human PAR3 amino acid
 CC sequence (see AAW51406) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC compounds for their ability to act as agonists or antagonists of
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,
 CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.

XX SQ Sequence 369 AA;

Query Match 90.8%; Score 1940; DB 19; Length 369;
 Best Local Similarity 100.0%; Pred. No. 2e-199;
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 MKLILVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSPNGQPNTFEEPLSDIEGWTG 76
 Db 1 MKLILVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSPNGQPNTFEEPLSDIEGWTG 60

Qy 77 ATTITKAECPEDSISTLHVNNATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLSL 136
 Db 61 ATTITKAECPEDSISTLHVNNATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLSL 120

Qy 137 RTKSISLIVFHTNLAIADLLFCVTLPPKIAVHLGNWNVGEVMCRITTVFVGNMYCAI 196
 Db 121 RTKSISLIVFHTNLAIADLLFCVTLPPKIAVHLGNWNVGEVMCRITTVFVGNMYCAI 180

Qy 197 LILTCMGINRYLATAPFTYOKLPKRSFSLMCGIVWVFLYMLPFVILKQEVHLVHSE 256
 Db 181 LILTCMGINRYLATAPFTYOKLPKRSFSLMCGIVWVFLYMLPFVILKQEVHLVHSE 240

Qy 257 ITTCHDVVDACESPSSRFYFVSLAFEGFLIPVILIFCYTTLIHKLKSKDRILWLYIK 316
 Db 241 ITTCHDVVDACESPSSRFYFVSLAFEGFLIPVILIFCYTTLIHKLKSKDRILWLYIK 300

Qy 317 AVLLILVIFTCFAPTNIILVIHANYHHNTDSLYFMYLIALCLGSLNSCLDPFLYFVM 376
 Db 301 AVLLILVIFTCFAPTNIILVIHANYHHNTDSLYFMYLIALCLGSLNSCLDPFLYFVM 360

Qy 377 SKVVDQLNP 385
 Db 361 SKVVDQLNP 369

RESULT 2

ID AAW51406
 AA AAW51406 standard; Protein; 374 AA.

AC AAW51406;

DT 12-OCT-1998 (first entry)

DE Human protease-activated receptor 3 (PAR3).

XX Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis.

XX Homo sapiens.
 OS Key Location/Qualifiers
 XX Modified-site 25..27 /note= "Asn is N-glycosylated"
 FT Cleavage-site 38..39 /note= "thrombin cleavage site"
 FT Peptide 48..51 /note= "hirudin-like sequence"
 FT Modified-site 82..84 /note= "Asn is N-glycosylated"
 FT Domain 95..117 /label= TM1 /note= "transmembrane domain 1"
 FT Domain 126..149 /label= TM2 /note= "transmembrane domain 2"
 FT Domain 168..191 /label= TM3 /note= "transmembrane domain 3"
 FT Domain 207..231 /label= TM4 /note= "transmembrane domain 4"
 FT Domain 261..286 /label= TM5 /note= "transmembrane domain 5"
 FT Domain 301..323 /label= TM6 /note= "transmembrane domain 6"
 FT Modified-site 331..333 /note= "Asn is N-glycosylated"
 FT Domain 335..360 /label= TM7 /note= "transmembrane domain 7"
 FT FT
 FT WO9818456-A1.
 XX 07-MAY-1998.
 XX 29-OCT-1997; 97WO-US19732.
 XX 30-OCT-1996; 96US-0742440.
 XX (REGC) UNIV CALIFORNIA.
 XX Connolly A, Coughlin SR, Ishihara H;
 XX WPI; 1998-271905/24.
 XX N-PSDB; AAV07374.
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX Claim 3; Page 41-42; 74pp; English.
 XX This polypeptide comprises human protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372), and shows homology to human PAR1
 CC and PAR2 (see AAW51407 and AAW51408). The mouse PAR3 amino acid
 CC sequence (see AAW51405) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC compounds for their ability to act as agonists or antagonists of
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,

CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.

XX SQ Sequence 374 AA;
 Query Match 66.7%; Score 1425.5; DB 19; Length 374;
 Best Local Similarity 71.6%; Pred. No. 3.5e-144;
 Matches 260; Conservative 44; Mismatches 58; Indels 1; Gaps 1;
 QY 17 MKILILVAAGLLFLPVTVCOSGI-NVSDNSAKPTLTIKSFNGGPQNTFFBFPPLSDIEGWT 75
 DB 1 MKALIPAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPNSFEFFPSALEGWT 60
 QY 76 GATTIKABCPEDSISLHNATIGYRLSLSSTQVIPAIIYLLFVGVGPNVITLWKL 135
 DB 61 GATITVKIKPEESASHLVKNATMGYLTSLSTKLIPAIYLLVGVGVPANAVTLWMLF 120
 QY 136 LRKTSISLVFHTNLADLLFCVTLFPKIAHYHLNGNNWVGEVCMCRITTVVFGNMYCA 195
 DB 121 FRRSICITTVFYTNLADELFCVTLFPKIAHYHLNGNNWVGEVLCRATTIVIFGNMYS 180
 QY 196 ILILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIYVNMVFLYMLPFVILKOEYHLVHS 255
 DB 181 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLWATVFLYMLPFVILKOEYILVQP 240
 QY 256 EITTCNDVDACSPSSFRFYFVSLAPFGFLIPFVIIIFCYTTLIHKLKSKDRIMLYI 315
 DB 241 DIITCHDVHNTCSSPFLYFISLAPFGFLIPFVIIIFCYTTLIHKLKSKDRIMLYI 300
 QY 316 KAVLLIIVITICFAPNTIILVIHHANYHYHNTDSLYPMYLIALCLGSLNSCLDPFLYFV 375
 DB 301 KASLLIILVITICFAPSNIILIIHHANYHYHNTDGLYFIYLIALCLGSLNSCLDPFLYF 360
 QY 376 MSK 378
 DB 361 MSK 363

RESULT 3

AA066922
 ID AAR66922 standard; Protein; 399 AA.
 XX AC AAR66922;
 XX DT 22-AUG-1995 (first entry)
 XX DE Murine C140 receptor deduced from cDNA.
 XX KW G-protein-coupled receptor; G-protein; C140 receptor.
 XX OS Mus musculus.
 XX PN WO9503318-A.
 XX PD 02-FEB-1995.
 XX PF 26-JUL-1994; 94WO-US08536.
 XX PR 26-JUL-1993; 93US-0057938.
 XX PA (CORT-) COR THERAPEUTICS.
 XX PI Scarborough RM, Sundelin J;
 XX DR WPI; 1995-075182/10.
 XX DR N-PSDB; AAQ84559.
 XX PT New DNA encoding recombinant C140 receptor - and novel agonists
 XX PT and antagonists and specific antibodies with therapeutic and
 XX PT diagnostic applications.
 XX PS Example; Fig 10; 57pp; English.
 XX CC A cDNA library from a mouse stomach was constructed in lambda gt10

CC and screened with a probe encompassing the C140 genomic clone (see
 CC AAQ84557). A single phage clone was isolated and cut with EcoRI. The
 CC insert was cloned in pBluescript and pSG5 and sequenced. The
 CC complete nt sequence and deduced AA sequence is given in AAQ84559 &
 CC AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
 CC end. The 5' end of the apparent coding region differs from the 5'
 CC end of the ORF of genomic DNA; it is believed that the 5' end of
 CC the cDNA sequence is correct.

XX SQ Sequence 399 AA;

Query Match 27.2%; Score 581.5; DB 16; Length 399;
 Best Local Similarity 37.6%; Pred. No. 1.6e-53;
 Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
 QY 95 VNNATTGYLRSSLSSTQVIPAIIYLLFVGVGPNVITLWKLSTRT-KSISLVIFHTNLIA 153
 DB 63 IDEFSASITGKLTITVFLPVIIIVFVIGLPSNGMALWIFLFTKKHPAVIYMANLALA 122
 QY 154 DLLFCVTLFPKIAHYHLNGNNWVGEVCMCRITTVVFGNMYCAIILTCMGINRYLATAHP 213
 DB 123 DLLSVIWFPLKISYHLGNWVYGEALCKVLIGFFYGNMVCISILFWTCLSVQRYWVIVNP 182
 QY 214 FTYQKLPKRSFSLMCGIYVNMVFLYMLPFVILKOEYHLVHSIITTCNDVDACSPSSF 273
 DB 183 MGHPR-KKANIAGVSLAIWLIFLVTIPLYVMKQTIYIPALNITTCNDVLP-EVLVGD 240
 QY 274 REYVYVSLAFAFGFLIPFVIIIFCYTTLI-----HKLKSKDRIMLYIKAVLLILV 323
 DB 241 MFNYFUSLAIGVFLPALLTASAYLVIMKTLRSSAMDEHSEKQRQA----ILIIITVLA 296
 QY 324 IFTICFAPNTIILVIHHANYHYHNTDSLYPMYLIALCLGSLNSCLDPFLYFVMSK 378
 DB 297 MYFICFAPSNIILVHYFLIKTQROSHVYALIVLALCLSLNSCIDPFVYFVYSK 351

RESULT 4

AAW01954
 ID AAW01954 standard; Protein; 399 AA.
 XX AC AAW01954;
 XX DT 02-APR-1997 (first entry)
 XX DE Murine C140 receptor.
 XX KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
 XX KW antagonist; hypertension; hypotension; blood pressure.
 XX OS Mus sp.
 XX FH Key Location/Qualifiers
 XX FT Peptide 1..31 /note= "the signal peptide differs from that
 XX FT encoded by a genomic DNA sequence for
 XX FT this receptor (see AAW01952), the signal
 XX FT sequence given here is believed to be
 XX FT the correct sequence"
 XX FT Protein 32..399 /note= "mature protein"
 XX PN WO9623225-A1.
 XX PD 01-AUG-1996.
 XX PF 25-JAN-1996; 96WO-US01179.
 XX PR 25-JAN-1995; 95US-0390301.
 XX PA (CORT-) COR THERAPEUTICS INC.
 XX PI Scarborough RM, Sundelin J;
 XX CC

DR WPI; 1996-362813/36.
 DR N-PSDB; AAT32038.
 XX Vector for expression C140 cell surface receptor in host cell
 PT useful to identify C140 agonist and antagonists, which are
 PT antihypertensives and elevators of blood pressure, respectively
 XX
 PS Example 4; Fig 10A-B; 60pp; English.
 XX
 CC AAW01954 represents the murine C140 receptor (C140R). DNA encoding C140R
 CC may be engineered so as to allow the recombinant expression of C140R in
 CC a suitable host cell, i.e. by removing the native expression-control
 CC sequences and replacing them with control sequences operable in the
 CC host. Such a recombinant receptor can be expressed on the surface of
 CC oocytes, this provides a good assay system for identifying
 CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
 CC receptor and a member of the "seven-pass" transmembrane receptor
 CC superfamily (peptide chain of the receptor passes through the cell
 CC membrane seven times, producing seven transmembrane regions within the
 CC receptor molecule). The C140 receptor is involved in controlling blood
 CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
 CC signalling from this receptor, resulting in an increase in blood pressure
 CC and are therefore useful in pharmaceuticals for the treatment of
 CC hypotension (low blood pressure). Conversely agonists (see
 CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
 CC of hypertension (high blood pressure).
 XX
 SQ Sequence 399 AA;

Query Match 27.2%; Score 581.5; DB 17; Length 399;
 Best Local Similarity 37.6%; Pred. No. 1.6e-53;
 Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
 QY 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIEHTNLAIA 153
 Db 63 IDEFSASILTGLTTFVPIVYIIIVFVIGLPSNGMALWIFLFTKXKHPAVIYMANLALA 122
 QY 154 DLLFCVTLFPFKIAYHLNGNNVFGVEMCRITTVVFGYNNMYCAILILTCMGINRYLATAHP 213
 Db 123 DLLSVIWFPLKISYHLGNNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182
 QY 214 FTYQKLPKRSFSLMCGIIVMVVFLYMLPFVILKQEVHLVHSITTCDDVDVADACESPSS 273
 Db 183 MGHPR-KKANIAVGSALWLLIFLVTIPIYVNMKQTIYIPALNITTCDDVLPB-EVLVGD 240
 QY 274 RFYFVSLAFGLPIPFVILFICVTTLI-----HKLKSKDRILWLGIVKAVLLILV 323
 Db 241 MFNYFLSLAIGVFLFPALLTASAYVLMIKLRSAMDEHSEKKRQRA----IRLIITVLA 296
 QY 324 IFTICFAPTNILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
 Db 297 MYFICFAPSNLLLVVHFVFLKIQRSQSHVYALYVALCLSLNSCLDPPFLYFVMSK 351

RESULT 5
 AAR66920
 ID AAR66920 standard; Protein; 395 AA.
 XX
 AC AAR66920;
 XX
 DT 22-AUG-1995 (first entry)
 XX
 DE Murine C140 receptor.
 XX
 KW G-protein-coupled receptor; G-protein; C140 receptor.
 XX
 OS Mus musculus.
 XX
 FH Key
 FT Protein 1..27 Location/Qualifiers
 FT /label= tentative signal sequence
 FT Region 78..100
 FT /label= transmembrane I

FT Region 108..128
 FT /label= transmembrane II
 FT Region 168..191
 FT /label= transmembrane III
 FT Region 168..191
 FT /label= transmembrane IV
 FT Region 204..224
 FT /label= transmembrane V
 FT Region 266..286
 FT /label= transmembrane VI
 FT Region 304..325
 FT /label= transmembrane VII
 FT Modified-site 29
 FT /label= Asn linked glycosylation site
 FT Cleavage-site 34..35
 FT /label= protease receptor cleavage
 FT Modified-site 200
 FT /label= Asn linked glycosylation site
 FT
 WO9503318-A.
 XX
 PD 02-FEB-1995.
 XX
 PF 26-JUL-1994; 94WO-US08536.
 XX
 PR 26-JUL-1993; 93US-0097938.
 XX
 PA (CORT-) COR THERAPEUTICS.
 XX
 PI Scarborough RM, Sundelin J;
 XX
 DR WPI; 1995-075182/10.
 DR N-PSDB; AAR64557.
 XX
 XX
 PT New DNA encoding recombinant C140 receptor - and novel agonists
 PT and antagonists and specific antibodies with therapeutic and
 PT diagnostic applications.
 PT
 PS Disclosure; Fig 1; 57pp; English.
 XX
 CC A mouse cosmid genomic library (obtd. from Dr R.A. Wetsel, Washington
 CC Univ. School of Medicine, St Louis, Missouri) was screened with two
 CC 32p-labeled oligos corresp. to bp 190-249 and 742-801 of the bovine
 CC substance K receptor cDNA. In one of the clones isolated (C140) the
 CC hybridising region was localised to a 3.7 kb pSTI fragment. This
 CC fragment was subcloned into pBluescript vector. The hybridising and
 CC adjacent regions were sequenced. The nt sequence and the deduced AA
 CC sequence are given in AAR64557 & AAR69920 respectively.
 XX
 SQ Sequence 395 AA;
 Query Match 26.9%; Score 574.5; DB 16; Length 395;
 Best Local Similarity 37.3%; Pred. No. 8.8e-53;
 Matches 110; Conservative 66; Mismatches 102; Indels 17; Gaps 5;
 QY 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIEHTNLAIA 153
 Db 59 IDEFSASILTGLTTFVPIVYIIIVFVIGLPSNGMALWIFLFTKXKHPAVIYMANLALA 118
 QY 154 DLLFCVTLFPFKIAYHLNGNNVFGVEMCRITTVVFGYNNMYCAILILTCMGINRYLATAHP 213
 Db 119 DLLSVIWFPLKISYHLGNNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
 QY 214 FTYQKLPKRSFSLMCGIIVMVVFLYMLPFVILKQEVHLVHSITTCDDVDVADACESPSS 273
 Db 179 MGHPR-KKANIAVGSALWLLIFLVTIPIYVNMKQTIYIPALNITTCDDVLPB-EVLVGD 236
 QY 274 RFYFVSLAFGLPIPFVILFICVTTLIHLKLS-----KDRILWLGIVKAVLLILV 323
 Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKLRSAMDEHSEKKRQRA----IRLIITVLA 292
 QY 324 IFTICFAPTNILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
 XX

FT	Region	11..132	
FT	FT	/label= transmembrane II	
FT	Region	150..174	
FT	FT	/label= transmembrane III	
FT	Region	191..212	
FT	FT	/label= transmembrane IV	
FT	Region	245..267	
FT	FT	/label= transmembrane V	
FT	Region	289..309	
FT	FT	/label= transmembrane VI	
FT	Region	327..348	
FT	FT	/label= transmembrane VII	
XX			
PN	W09503318-A.		

XX	02-FEB-1995.	
PD		
XX		
XX	26-JUL-1994; 94WO-US08536.	
PF		
XX		
XX	26-JUL-1993; 93US-0097938.	
PR		
XX	(CORT-) COR THERAPEUTICS.	
PA		
XX		
XX	Scarborough RM, Sundelin J;	
PI		
XX	WPI; 1995-075182/10.	
XX	N-PSDB; AAQ84558.	
DR		
XX	New DNA encoding recombinant Cl40 receptor - and novel agonists	
PT	and antagonists and specific antibodies with therapeutic and	
PT	diagnostic applications.	
XX		
PS	Disclosure; Fig 2; 57pp; English.	
XX		
XX	The availability of genomic DNA encoding the mouse protease	
CC	Cl40 receptor (see Q84557) permitted the retrieval of the corresp.	
CC	human gene. A human genomic library cloned in the vector EMBL3 was	
CC	screened using the entire coding region of the murine clone as a	
CC	probe. The recovered human gene including the DNA sequence and the	
CC	deduced AA sequence are shown in Q84558 & R65921. Subsequent	
CC	experiments indicated that the human Cl40 gene is located in the	
CC	same region of the long arm of chromosome number 5 (5q12-5q13)	
CC	as has been reported for the human thrombin receptor gene.	
XX		
XX	Sequence 398 AA;	
SQ		

Best Local Similarity	33.5%;	72;	Mismatches	128;	Indels	28;	Gaps	7
Matches	115;	Conservative						
QY	49	TLTIKSFNGSPONTFEEBFLPSLDIEGW---	TGATTITIKAECPEDSITLTHVNNAITGYLRS	105				
DB	21	TLVFLSCTGNRSKGRSLIGKVDGSHVTKGVTE-----	TVFSVDEFSASVLTG	72				
QY	106	SLSTQVPIPAIYILLFVVGVSPSNTIWLKLSLRT-KSISLVI	FHTLAIADLLFCVTLDPFK	164				
DB	73	KLTTFELPIVYTVFVVGFLSNGMALWVFLFRTKKH	PAIYMANLADLLSVIFWPLK	132				
QY	165	IAYHLNGNNVFGHEWCRITTVFYGNMYCAIILTCMGINRYLA	TAHPFTYQKLPKRSF	224				
DB	133	IAYHIGNNNIYGEALCNVLIGFFYGNMYCSILFMTCLS	VQRYWYIVTNPMGHSR-KKANI	191				
QY	225	SLLMCGIYVWVFLMYLPPFVILKQYHLVHSEITTC	HDVVDVACEPSFSF----	280				
DB	192	AIGISLAIWLILLLVPLVXVVKQTTFIPALNLTTC	HDVL-----	246				
QY	281	LAFFGFLPIPVIIIFCYTTLLIHLKS-----	KDRIWLGVYKAVLLILVIFTCFAPNTI	334				
DB	247	LAIGVFLFPAFLTASAVLMIRLMSRWSENSEKKRKA	IKLIVTLAMYLICTPSPNL	306				
QY	335	ILVTHIHHANYHNDSLYFWYLIACLGSLNSCLDPFL	YFVMS	377				
DB	307	LLVHYFLIKSQGSHVYALYIVALCLSTNSCIDP	YFVYFVS	349				

RESULT 9
AAW01953
ID AAW01953 standard; Protein; 398 AA.
XX
AC AAW01953;
XX
DT 01-APR-1997 (first entry)
XX
XX Human C140 receptor, with putative signal sequence.
DE
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /note= "putative signal peptide, differs from
FT signal peptide encoded by a cDNA clone of
FT this receptor (see AAW01955), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT
FT Protein 28..398
FT /note= "mature protein"
FT Modified-site 31
FT /note= "potential Asn-linked glycosylation site"
FT
FT Cleavage-site 37..38
FT /note= "putative protease receptor cleavage site"
FT Region 81..103
FT /note= "transmembrane region I"
FT Region 111..132
FT /note= "transmembrane region II"
FT Region 151..174
FT /note= "transmembrane region III"
FT Region 191..212
FT /note= "transmembrane region IV"
FT Modified-site 223
FT /note= "potential Asn-linked glycosylation site"
FT Region 245..267
FT /note= "transmembrane region V"
FT Region 289..309
FT /note= "transmembrane region VI"
FT Region 327..348
FT /note= "transmembrane region VII"
XX
PN W09623225-A1.
XX
XX
PD 01-AUG-1996.
XX
XX 25-JAN-1996; 96WO-US01179.
XX
XX 25-JAN-1995; 95US-0390301.
XX
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Scarborough RM, Sundelin J;
XX
XX WPI; 1996-362813/36.
DR N-PSDB; AAT32037.
XX
XX Vector for expression C140 cell surface receptor in host cell
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
PT
XX Example 2; Fig 2A-B; 60pp; English.
XX
XX AAW01953 represents the human C140 receptor (C140R), including a
CC putative signal peptide (see features table). DNA encoding C140R may be
CC engineered so as to allow the recombinant expression of C140R in a
CC suitable host cell, i.e. by removing the native expression-control
CC sequences and replacing them with control sequences operable in the

CC host. Such a recombinant receptor can be expressed on the surface of
CC oocytes, this provides a good assay system for identifying
CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
CC receptor and a member of the "seven-pass" transmembrane receptor
CC superfamily (peptide chain of the receptor passes through the cell
CC membrane seven times, producing seven transmembrane regions within the
CC receptor molecule). The C140 receptor is involved in controlling blood
CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
CC signalling from this receptor, resulting in an increase in blood pressure
CC and are therefore useful in pharmaceuticals for the treatment of
CC hypotension (low blood pressure). Conversely agonists (see
CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
CC of hypertension (high blood pressure).
XX
XX Sequence 398 AA;
SQ
Query Match 26.4%; Score 563; DB 17; Length 398;
Best Local Similarity 33.5%; Pred. No. 1.5e-51;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;
QY 49 TLTIKSFNGGPONTFEFPLSDIEGW--TGATTIKAECPEDSISTLHVNNATIGYLR 105
DB 21 TLVFLSCTGTNRSSKGRSLIGKVDGTSHTVGKVTVE-----TVFSVDEFSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVGVPSNIVTLWKLRLT-KSISLVFIHTNLAIAADLLFCVTLPFK 164
DB 73 KLTVFLPIVYITIVFVVGPSNGMALWFLFRTKKKHPAVYMANLADLLSVIWFPLK 132
QY 165 IAYHLNGNNVFGVEMCRITTVFYGNMYCAILLTCMGINRYLATAHPTYQKLPKRSF 224
DB 133 IAYHNGNNIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVWINPMGHSR-KKANI 191
QY 225 SLLMCGIVVMVFLYMLPFVILKQEVHLVHSEITTCDDVVDACESPSSF----RFYFVS 280
DB 192 AIGISLAIWLLILLVTIPLYVVKQITFIPALMITTCHDVL-----PEQLLVGDMFNYFLS 246
QY 281 LAFFGFLIPFVILFCYTTLIHLKLS-----KDRIMLVYKAVLLILVIFTICFAPTNI 334
DB 247 LAIGVFLPFAFLTASAYVLMIRLSSAMDENSEKKRRAIKLIVTVLAMYLICFTPSNL 306
QY 335 ILVVIHHANYYYHNTDSLYFWYLIALLCLGSLNSCLDPLFYFVMS 377
DB 307 LLVWHYFLIKSQGSHVYALYIVALCLCLTSLNSCIDPFYFVVS 349
RESULT 10
AAW51408
ID AAW51408 standard; Protein; 394 AA.
XX
XX AAW51408;
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 2 (PAR2).
XX
KW Protease-activated receptor 2; PAR2; PAR3; thrombin receptor;
XX human.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Cleavage-site 36..37
FT /note= "thrombin cleavage site"
XX
PN W09818456-A1.
XX
XX 07-MAY-1998.
XX
XX 29-OCT-1997; 97WO-US19732.
XX
XX 30-OCT-1996; 96US-0742440.
XX
XX (REGC) UNIV CALIFORNIA.
PA

XX Connolly A, Coughlin SR, Ishihara H;
XX WPI; 1998-271905/24.
XX
XX DNA encoding protease-activated receptor 3 - for detection of
XX specific agonists and antagonists, potentially useful for treating
XX e.g. thrombosis, atherosclerosis, inflammation etc.
XX
XX Example 1; Page 43-44; 74pp; English.
XX
XX This polypeptide comprises human protease-activated receptor 2
XX (PAR2). The physiological activator of PAR2 remains unknown;
XX it is not activated by thrombin. The invention relates to novel
XX mouse and human PAR3 (see AAW51405-06) that show homology to PAR2 and
XX which are specific receptors for thrombin. They can be used to
XX screen for specific agonists and antagonists of thrombin useful
XX e.g. for treating atherosclerosis, thrombosis and inflammation.
XX
XX Sequence 394 AA;

Query Match 25.9%; Score 553.5; DB 19; Length 394;
Best Local Similarity 32.2%; Pred. No. 1.6e-50;
Matches 118; Conservative 73; Mismatches 136; Indels 39; Gaps 8;

QY 22 LVAAGLLFLPVTVCOSGINVSDNSAKPTLTIKSPNGGPQNTFEFFPLSDIEGWTGATTI 81
DB 9 LLGAAILLAAASLSCSTIOGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSQVIPAIIYLLFVVGPSNIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDEFSASVLTGKLTTFVLPVIVTVFVVGPSNGMALWVFLFRKK 107
QY 141 ISLVIFHTNLAIADLLFCVTLPFKIAIHLNNGNWFGEVNCRIITVVFYGNMYCAILILT 200
DB 108 HPAVIYMANLALADLLSVIWFPLKIAIYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFSLMCGIWMVVMVFLYMLPFVILKQEHVHSEITTC 260
DB 168 CLSVQRWVIVNPMGHSR-KKANAIGISLAIWLLTLVTVPLVYVVKQTFIPALNITTC 226
QY 261 HDVVD---ACESPSSFRFYVSLAFGLIPFVIIIFCYTTLIHKLK-----KDRIW 311
DB 227 HDVLPQELLVGDP-----FLSLAIGVFLPAPLTASAVVLMIRLSSAMENSEKKR 279
QY 312 LGYIKAVLLILVIFTCFAPTNILVIHANYHHNTDSLYFMYLIALCLGSLNSCLDPF 371
DB 280 KRAIKLIVTLAMYLICTPNSLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDPF 339
QY 372 LYFVMS 377
DB 340 VYFVVS 345

RESULT 11
AAW66923
ID AAR66923 standard; Protein; 397 AA.
XX
XX AAR66923;

XX 22-AUG-1995 (first entry)
XX Human C140 receptor encoded by cDNA.
XX G-protein-coupled receptor; G-protein; C140 receptor.
XX Homo sapiens.

XX WO9503318-A.
XX 02-FEB-1995.
XX 26-JUL-1994; 94WO-US08536.

XX 26-JUL-1993; 93US-0097938.
XX (CORT-) COR THERAPEUTICS.
XX
XX Scarborough RM, Sundelin J;
XX WPI; 1995-075182/10.
XX N-PSDB; AAQ84560.
XX
XX New DNA encoding recombinant C140 receptor - and novel agonists
XX and antagonists and specific antibodies with therapeutic and
XX diagnostic applications.
XX
XX Example; Fig 11; 57pp; English.
XX
XX A human intestinal tumour cDNA library was subjected to PCR using
XX primers designed from the genomic clone (see AAQ84558) and the
XX amplified fragment was cloned in pSG5 and sequenced. There are
XX four AA differences between the cDNA encoded sequence and that
XX encoded by the genomic DNA. The genomic DNA sequence and deduced
XX AA sequence are given in AAQ84560 & AAR66923.
XX
XX Sequence 397 AA;

Query Match 25.8%; Score 551; DB 16; Length 397;
Best Local Similarity 32.2%; Pred. No. 3e-50;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;

QY 22 LVAAGLLFLPVTVCOSGINVSDNSAKPTLTIKSFNGGPQNTFEFFPLSDIEGWTGATTI 81
DB 9 LLGAAILLAAASLSCSTIOGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSQVIPAIIYLLFVVGPSNIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDEFSASVLAGKLTTFVLPVIVTVFVVGPSNGMALWVFLFRKK 107
QY 141 ISLVIFHTNLAIADLLFCVTLPFKIAIHLNNGNWFGEVNCRIITVVFYGNMYCAILILT 200
DB 108 HPAVIYMANLALADLLSVIWFPLKIAIYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFSLMCGIWMVVMVFLYMLPFVILKQEHVHSEITTC 260
DB 168 CLSVQRWVIVNPMGHSR-KKANAIGISLAIWLLTLVTVPLVYVVKQTFIPALNITTC 226
QY 261 HDVVDACESPSSFRFYVSLAFGLIPFVIIIFCYTTLIHKLK-----KDR 310
DB 227 HDVL-----PEQLLVGDMFNYFLSLAIGVFLPAPLTASAVVLMIRLSSAMENSEK 281
QY 311 WLGYIKAVLLILVIFTCFAPTNILVIHANYHHNTDSLYFMYLIALCLGSLNSCLDP 370
DB 282 KRAIKLIVTLVGLMYLICTPNSLLLVVHYFLIKSQGQSHVYALYIVALCLSTLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 12
AAW01955
ID AAW01955 standard; Protein; 397 AA.
XX
XX AAW01955;

XX 02-APR-1997 (first entry)
XX Human C140 receptor.
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure.
XX Homo sapiens.

Db 121 IMAIVFILKMKVKK-PAVVVMLHLATADVLSVLPFKISYFSGSDMQFSELRCFVT 179
 QY 186 VVFYGNMYCAIILTCMGINRYLATAHP---FTYQKLPKPSFSLMCGIYVWVFLYMLP 242
 Db 180 AAFYCNMYASILEMTVISIDRFVAVVPMQSLWRTLGASFT---CLAIWALAIAGVWP 236
 QY 243 FVILKQEVHLVHSEITTCDDVDVADACESPSGRFYFVSLAFFGFLIPEFVILIFCYTTLIH 302
 Db 237 LVLKEQTIQVPLNITTCDDVDVADACESPSGRFYFVSLAFFGFLIPEFVILIFCYTTLIH 295
 QY 303 KLKS-----KDRILWGLYIKAVLL---ILVFTICFAPTNILVILVHANYHH--NTD 349
 Db 296 CLSSSAVANRSKKR-----ALFLSAAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
 QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 Db 348 AAFAYALLCCVSSISSCIDPLIYYAS 375

RESULT 16

AAW51407

ID AAW51407 standard; Protein; 425 AA.

AC AAW51407;

DT 12-OCT-1998. (first entry)

DE Human protease-activated receptor 1 (PAR1).

KW Protease-activated receptor 1; PAR1; PAR3; thrombin receptor;

KW human.

OS Homo sapiens.

FH Key

FT Binding-site

FT /note= "thrombin binding site"

FT Cleavage-site

FT /note= "thrombin cleavage site"

XX WO9818456-A1.

XX 07-MAY-1998.

XX 29-OCT-1997; 97WO-US19732.

XX 30-OCT-1996; 96US-0742440.

XX (REGC) UNIV CALIFORNIA.

XX Connolly A, Coughlin SR, Ishihara H;

XX WPI; 1998-271905/24.

XX DNA encoding protease-activated receptor 3 - for detection of

XX specific agonists and antagonists, potentially useful for treating

XX e.g. thrombosis, atherosclerosis, inflammation etc.

XX Example 1; Page 42-43; 74pp; English.

XX This polypeptide comprises human protease-activated receptor 1

XX (PAR1), a receptor that mediates thrombin signalling. The

XX invention relates to novel mouse and human PAR3 (see AAW51405-06)

XX that show homology to PAR1 and which are specific receptors for

XX thrombin. They can be used to screen for specific agonists and

XX antagonists of thrombin useful e.g. for treating atherosclerosis,

XX thrombosis and inflammation.

XX Sequence 425 AA;

SQ Query Match

Best Local Similarity 23.8%; Score 509; DB 19; Length 425;

Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSF-NGGPONTREERPLSDI 71
 Db 6 LLLVAACF-----SLCGPLSARTARRPESKATNATLDPKSLRNPNKDYEPFWEDEE 60
 QY 72 EGMTGAT---TTKAECPEDSISTLHVNNATIGYLRSLSTQVIPAIIYILLFVVGVPSN 127
 Db 61 KNEISGLETVRLVYSINKSSPLOKQLPAFISEDAGYLTSSMLTLFVPSVVTGVFVLSPLN 120
 QY 128 I---VTLWKLSTKTSISLVIPTNLAIAIDLFCVTLPFKIAIHLNNGNNVFGVEMCRIT 185
 Db 121 IMAIVFILKMKVKK-PAVVVMLHLATADVLSVLPFKISYFSGSDMQFSELRCFVT 179
 QY 186 VVFYGNMYCAIILTCMGINRYLATAHP---FTYQKLPKPSFSLMCGIYVWVFLYMLP 242
 Db 180 AAFYCNMYASILEMTVISIDRFVAVVPMQSLWRTLGASFT---CLAIWALAIAGVWP 236
 QY 243 FVILKQEVHLVHSEITTCDDVDVADACESPSGRFYFVSLAFFGFLIPEFVILIFCYTTLIH 302
 Db 237 LVLKEQTIQVPLNITTCDDVDVADACESPSGRFYFVSLAFFGFLIPEFVILIFCYTTLIH 295
 QY 303 KLKS-----KDRILWGLYIKAVLL---ILVFTICFAPTNILVILVHANYHH--NTD 349
 Db 296 CLSSSAVANRSKKR-----ALFLSAAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
 QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 Db 348 AAFAYALLCCVSSISSCIDPLIYYAS 375

RESULT 17

AA49570

ID AAY49570 standard; Protein; 425 AA.

AC AAY49570;

DT 13-JAN-2000 (first entry)

DE Human thrombin receptor protein sequence.

KW Human; coding sequence polymorphism; vascular pathology gene;

KW polymorphic site; phenotype correlation; forensic; paternity testing;

KW medicine; genetic analysis; vascular disease.

OS Homo sapiens.

XX WO950454-A2.

XX 07-OCT-1999.

XX 26-MAR-1999; 99WO-US06473.

XX 01-APR-1998; 98US-0054272.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

XX WPI; 1999-620066/53.

XX N-PSDB; AAZ32191.

XX Determination of polymorphisms in genes, especially those identifying

XX predisposition to vascular disease

XX Disclosure; Fig 35; 134pp; English.

XX AAZ32159 to AAZ32194 represent reference alleles for specifically

XX claimed nucleic acid sequences from the present invention which comprise

XX polymorphic sites as given in a table in the specification, selected

XX from 92 single nucleotide polymorphisms in which the nucleotide at the

XX polymorphic site is different from a nucleotide at the same site in a

XX reference allele. The nucleic acids, and primers and probes, are used to

XX identify polymorphisms, which may predispose an individual to disease.

CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AAY49550 to AAY49573 represent the proteins which correspond
CC to some of the reference alleles.
XX
SQ Sequence 425 AA;

Query Match 23.8%; Score 509; DB 20; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.1e-45;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

Qy 20 LILVAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSP-NGCPONTPEEPFLSDI 71
Db 6 LLLVAACF-----SLCGLLSARTRARRPESKATNATLDRSFLRNPNKYEPPFWEDEE 60

Qy 72 EGMWGAT-----TTIKAECPEDSITLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPNS 127
Db 61 KNEGSLTEYRLVSLINKSPLOKQLPAFISEDAGSLTSSWLTLPVSPVYTGTVFVWSLPLN 120

Qy 128 I--VTLWKLSLRTKSIISLVIFHTNLAIADLLFCVTLPLPKIAYHLNNGNWFGEVVMCRITT 185
Db 121 IMAIWFLTKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYFSGSDWQFGSELCRFVT 179

Qy 186 VVFYGNMYCAIILITCMGIRYLATNP---FTYQKLPKPSFLMCGIIVWVFLWMLP 242
Db 180 AAFYCNMYASILLMTVISIDFLAVVYPMOSLWRTLGRASFT---CLAIWALAIAGVWP 236

Qy 243 FVILKQEVHLVSEIITCHDVVDACESSPSRFYFVSLAFPGFLIPFVLIIFCYTTLIH 302
Db 237 LVLKEQITQVGLMITTCHDVNLTEGYYIYFSAFV-FVFLIISTVCYVSIIR 295

Qy 303 KLKS-----KDRIMLGIXKAVLL---ILVIFTICFAPTNIIILVIHANYVH--NTD 349
Db 296 CLSSAVANRSKSR-----ALFLSAVFCIIICGPNVLLIAHY-SFLSHTSTTE 347

Qy 350 SLYFWYLIALLCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 18
AAG80697 standard; Protein; 425 AA.
XX AAG80697;
XX AC AAG80697;
XX DT 14-MAR-2002 (first entry)
XX DE Human thrombin protein.

KW G-protein coupled receptor; human; analgesic; antirheumatic; nootropic;
KW antiarthritic; neuroprotective; antiparkinsonian; hypotensive; cardiac;
KW antidepressant; antimigraine; cytotatic; antidiabetic; cardiac disorder;
KW muscular activity; immunosuppressive; vasotropic; gene therapy; cancer;
KW cellular processes modulation; intracellular signalling modulation;
KW proliferative disorder; central nervous system disorder; CNS; pain;
KW metabolic disorder; muscular disorder; rheumatoid arthritis; depression;
KW Alzheimer's disease; Parkinson's disease; autoimmune function disorder;
KW hypertension; neuropsychiatric disorder; learning disorder;
KW memory disorder; bipolar effective neurological disorder;
KW diabetes mellitus; autoimmune disorder; thrombin.

XX OS Homo sapiens.

FH Key Location/Qualifiers
FT Domain 102..132
FT /label= TM1
FT /note= "Transmembrane domain 1"
FT Domain 138..165
FT /label= TM2
FT /note= "Transmembrane domain 2"
FT Domain 176..200
FT /label= TM3

FT Domain /note= "Transmembrane domain 3"
FT 215..241
FT /label= TM4
FT /note= "Transmembrane domain 4"
FT 258..296
FT /label= TM5
FT /note= "Transmembrane domain 5"
FT 313..340
FT /label= TM6
FT /note= "Transmembrane domain 6"
FT 347..379
FT /label= TM7
FT /note= "Transmembrane domain 7"
XX WO200196400-A2.
XX 20-DEC-2001.
XX 18-JUN-2001; 2001WO-US19464.
XX 16-JUN-2000; 2000US-212331P.
XX 16-FEB-2001; 2001US-269758P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Glucksmann MA, Silos-Santiago I;
XX WPI; 2002-098057/13.
XX New isolated human G protein-coupled receptor polypeptide, 52871, for
XX diagnosing and treating proliferative, central nervous system, pain,
XX cardiac, metabolic or muscular disorders -
XX Disclosure; Page 106-107; 112pp; English.

CC This invention describes a novel isolated G protein-coupled receptor
CC (GPCR) polypeptide (I), designated 52871. The product of the invention
CC has analgesic, antirheumatic, antiarthritic, nootropic, neuroprotective,
CC antiparkinsonian, hypotensive, antidepressant, antimigraine, cytotatic,
CC antidiabetic, immunosuppressive, vasotropic, cardiac, muscular activity
CC and can be used for gene therapy, cellular processes modulation and
CC intracellular signalling modulation. (I), nucleic acid (II) encoding (I)
CC or an antibody derived from (I) is useful in screening assays, predictive
CC medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical
CC trials and pharmacogenomics), and in methods of treatment (e.g.,
CC therapeutic and prophylactic). (I) or (II) are useful for treating
CC 52871-associated disorders such as proliferative disorders, central
CC nervous system disorders, pain or pain disorders, cardiac disorders,
CC metabolic disorders or muscular disorders. (I) or (II) are useful for
CC treating pain disorders such as rheumatoid arthritis, CNS disorder such
CC as Alzheimer's disease, Parkinson's disease, autoimmune function disorders
CC such as hypertension, neuropsychiatric disorders such as depression,
CC learning or memory disorders, bipolar effective neurological disorders
CC such as migraine, cardiac-related disorders such as stenosis, hormonal
CC disorders such as diabetes mellitus, autoimmune disorders and cancer.
CC This sequence represents the human thrombin protein described in the
CC method of the invention.

XX Sequence 425 AA;

Query Match 23.8%; Score 509; DB 23; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.1e-45;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

Qy 20 LILVAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSP-NGCPONTPEEPFLSDI 71
Db 6 LLLVAACF-----SLCGLLSARTRARRPESKATNATLDRSFLRNPNKYEPPFWEDEE 60

Qy 72 EGMWGAT-----TTIKAECPEDSITLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPNS 127
Db 61 KNEGSLTEYRLVSLINKSPLOKQLPAFISEDAGSLTSSWLTLPVSPVYTGTVFVWSLPLN 120

Qy 128 I--VTLWKLSLRTKSIISLVIFHTNLAIADLLFCVTLPLPKIAYHLNNGNWFGEVVMCRITT 185

Db 121 IMAIVVFLKMKVKK-PAVVYMLHLATADVLVSVLPFKISYYPSGSDMQSGELCRFVT 179
Qy 186 VVFGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIIVVMVFLYMLP 242
Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
Qy 243 FVILKQEHVHSEITTCDDVDVADCESSPSFRFYFVSFLAFFGLIPFVLIIFCYTTLIH 302
Db 237 LVLKEQTIQVPGINITTCDDVDVADCESSPSFRFYFVSFLAFFGLIPFVLIIFCYTTLIH 295
Qy 303 KLKS-----KDRIMLGVYKAVLL---ILVFTICFAPTNIILVIHANYHH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
Qy 350 SLYFMYLIALLCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375
RESULT 19
AAE17032
ID AAE17032 standard; Protein; 425 AA.
AC AAE17032;
XX
DT 18-APR-2002 (first entry)
XX
DE Human thrombin protein.
XX
KW Human; G protein-coupled receptor; GPCR; SLGP; SLGP; cellular proliferation;
KW growth; differentiation; migration disorder; cancer; retinal; vasotropic;
KW optic disk neovascularisation; cytostatic; gene therapy; genetic disease;
KW tissue typing; forensic biology; antiarthritic; ischaemia; arthritis;
KW thrombin.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT 102..132
FT /label= TW1
FT /note= "Transmembrane domain"
FT 138..165
FT /label= TW2
FT /note= "Transmembrane domain"
FT 176..200
FT /label= TW3
FT /note= "Transmembrane domain"
FT 215..241
FT /label= TW4
FT /note= "Transmembrane domain"
FT 268..296
FT /label= TW5
FT /note= "Transmembrane domain"
FT 313..340
FT /label= TW6
FT /note= "Transmembrane domain"
FT 347..379
FT /label= TW7
FT /note= "Transmembrane domain"
FN WO200202602-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-US20751.
XX
PR 30-JUN-2000; 2000US-0608921.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
FI Tsai F;
XX

DR WPI; 2002-140083/18.
XX
PT New isolated G protein-coupled receptor nucleic acid molecule, SLGP,
PT useful for diagnosing and treating cellular proliferation, growth,
PT differentiation, or migration disorders e.g. cancer, arthritis and
XX myocardial ischemia
XX
PS Disclosure; Page 158-159; 182pp; English.
XX
CC The present invention relates to an isolated G protein-coupled receptor
CC (GPCR) nucleic acid molecule, SLGP which encodes a polypeptide. SLGP is
CC useful as target or therapeutic agent for diagnosis and treatment of
CC cellular proliferation, growth, differentiation, or migration disorders
CC (e.g., cancer, arthritis, retinal and optic disk neovascularisation, and
CC tissue ischaemia, such as myocardial ischaemia). SLGP is useful in
CC screening assays, detection assays (e.g., chromosomal mapping, tissue
CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,
CC prognostic assays, monitoring clinical trials and pharmacogenomics), and
CC in methods of treatment (e.g., therapeutic and prophylactic). SLGP is
CC useful as reagents or targets in assays applicable to treatment and
CC diagnosis of SLGP-mediated or related disorders. SLGP is useful in gene
CC therapy, to express SLGP protein, to detect SLGP mRNA or a genetic
CC alteration in a SLGP gene, and to modulate SLGP activity. SLGP is useful
CC to map their respective genes on a chromosome, and thus locate gene
CC regions associated with genetic disease, to identify an individual from a
CC minute biological sample (tissue typing), and to aid in forensic
CC identification of a biological sample. The present sequence is human
CC thrombin protein used in the invention.
XX
SQ Sequence 425 AA;
XX
Query Match 23.8%; Score 509; DB 23; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.1e-45;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
Qy 20 LILVAAGLLPLVPTVCQSGINVDNSAKP-----TLTIKSF-NGGPNQTEFFPLSDI 71
Db 6 LULVAACF-----SLCGPLLSARTRARRPESKATNATLDPFRLLRNPNDKYEPFWEDEE 60
Qy 72 ECGWTGAT---TTIKAECPEDSISTLHVNNATIGYLRSSLSQTQVIPAIVILLFVGVGPSN 127
Db 61 KNESGLTEYRLVLSINKSSPLQKLPAPISADASGYLTSSWLTFLVPSVYTGFWVSLPLN 120
Qy 128 I--VTLWKLRLTKSISLVIEHTNLAIADLLFCVTLPPKIAVHLNGNNVFGVEMCRIT 185
Db 121 IMAIVVFLKMKVKK-PAVVYMLHLATADVLVSVLPFKISYYPSGSDMQSGELCRFVT 179
Qy 186 VVFGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIIVVMVFLYMLP 242
Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
Qy 243 FVILKQEHVHSEITTCDDVDVADCESSPSFRFYFVSFLAFFGLIPFVLIIFCYTTLIH 302
Db 237 LVLKEQTIQVPGINITTCDDVDVADCESSPSFRFYFVSFLAFFGLIPFVLIIFCYTTLIH 295
Qy 303 KLKS-----KDRIMLGVYKAVLL---ILVFTICFAPTNIILVIHANYHH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
Qy 350 SLYFMYLIALLCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375
RESULT 20
AAE17032
ID AAY45035 standard; Protein; 426 AA.
XX
AC AAY45035;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human thrombin receptor, a member of protease activated receptor family.

CC attacks and strokes, and block inflammatory and proliferative responses
CC that occur in normal wound healing and variety of diseases including
CC atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
CC glomerulosclerosis.

XX SQ Sequence 385 AA;

Query Match 23.8%; Score 507.5; DB 20; Length 385;
Best Local Similarity 31.4%; Pred. No. 1.3e-45;
Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;

Qy 27 LLFLPVTCQSGINVSNSAKPTTIK--SPNGSQPTTFEFPDLSIEGWTGATTIKAE 84
Db 5 LLLWPLVL--GFSLSGQTQPSVYDESGTGGDDSTPSILPAP--RGYPGVQV----- 53

Qy 85 CPESISITLHVNNATIGYLRSSLSQVPIAIYILFVGVGPSNIVTWKLSLRTKYSISLV 144
Db 54 CANS-DTLEPLDSSRALLGWPTRLVPALYGLVVLVGLPANGALWVLATQAPRLPST 112

Qy 145 IFHTNLATADLLFCVTLPFKIAYHLNGNNVFGVMCRITTVFVYGNMYCAILLITCMGI 204
Db 113 MLNMLNATADLLALAPPRYHLRGORPFGAARLATAALYGHVGSVLLAAVSL 172

Qy 205 NRYLATARPFTYQKPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEIITCTHDV 264
Db 173 DRYLALVHLPRARALGRRLALGLCMAALMAALALPLTLQRTFRLARSDRVLC HDAL 232

Qy 265 DACSPSPFRYYVSVLAFGLFPFVILFVILFVILFVILFVILFVILFVILFVILFV 324
Db 233 -PLDAQSHWQPATCTALLGCFPLLAMLLCYGATLHTLAASGRYRGHALRTAVVLAS 291

Qy 325 FTICFAPTNILVTHANYVYHNTDSLVFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 292 AVAFVPSNLLLLHYDPSPSANGNLYGAVPSLSALSTLNSCVDPFIYYVVS 344

RESULT 22

AY50135
ID AAY50135 standard; Protein; 385 AA.

XX AC AAY50135;

DT 31-JAN-2000 (first entry)

XX DE Human protease-activated receptor PAR4.

XX KW Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;
KW platelet; proliferation; differentiation; mediation;
KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW growth factor; production.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Region 1..78 /note= "Extracellular N-terminal region"

FT Peptide 1..17 /note= "Signal peptide"

FT Cleavage-site 17..18 /note= "Cleaved by signal peptidase"

FT Protein 18..385 /note= "Mature non-activated human PAR4"

FT Cleavage-site 47..48 /note= "Cleaved by thrombin or trypsin to activate"

FT Protein 48..385 /note= "Activated human PAR4"

FT Region 48..53 /note= "Tethered hexapeptide activating ligand (AAY50140)" Modified-site 56

FT /note= "N-glycosylated"

FT Domain 79..102 /note= "Transmembrane domain 1"
FT Region 103..110 /note= "Intracellular loop 1"
FT Domain 111..132 /note= "Transmembrane domain 2"
FT Region 133..150 /note= "Extracellular loop 1"
FT Domain 151..172 /note= "Transmembrane domain 3"
FT Region 173..191 /note= "Intracellular loop 2"
FT Domain 192..213 /note= "Transmembrane domain 4"
FT Region 214..239 /note= "Extracellular loop 2, determines specificity for activating peptide"
FT Region 228..230 /note= "These three residues are conserved among PAR1-4"
FT Domain 240..263 /note= "Transmembrane domain 5"
FT Region 264..283 /note= "Intracellular loop 3"
FT Domain 284..305 /note= "Transmembrane domain 6"
FT Region 306..316 /note= "Extracellular loop 3"
FT Domain 317..343 /note= "Transmembrane domain 7"
FT Region 344..385 /note= "Intracellular C-terminal region"

XX WO950415-A2.

PN 07-OCT-1999.

XX 31-MAR-1999; 99WO-US07100.

XX 01-APR-1998; 98US-0053866.

XX (ZYMO) ZYMOGENETICS INC.

XX (UNIW) UNIV WASHINGTON.

XX Xu W, Presnell SR, Yee DP, Foster DC;

XX WPI; 1999-633640/54.

XX N-PSDB; AA232747, AA232748.

XX Novel protease activated receptor 4, useful for screening for
XX (antagonists for promoting the proliferation and/or differentiation of
XX platelets and in mediating inflammatory events -
XX Claim 8; Fig 1; 85pp; English.

XX This sequence represents a human protease-activated receptor, PAR4 (also
XX referred to as ZCHEMR2). The cDNA was identified from EST (expressed
XX sequence tag) sequences with homology to the three known protease-
XX activated receptors (PAR1, PAR2 and PAR3). Protease-activated receptors
XX (PARs) are a subfamily of G protein coupled receptors which are capable
XX of mediating cellular signalling in response to proteases (e.g.,
XX thrombin). They are characterised by a tethered peptide ligand at the
XX extracellular N-terminus that is generated by proteolysis. PAR4 is
XX activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates
XX a new N-terminus corresponding to the tethered ligand (a hexapeptide).
XX Agonists of PAR4 are useful for upregulating cellular or physiological
XX responses whereas antagonists are used to downregulate these
XX activities. The PAR4 protein is further useful for dissecting the
XX effects of thrombin or other activating proteases in the clotting
XX pathway from the effects of these proteases at the cellular level.
XX Agonists are specifically useful in promoting the proliferation
XX and/or differentiation of platelets, in mediating inflammatory events,
XX responses to vascular injury, chemotaxis or mitogenesis, and in
XX producing growth factors. Antagonists are useful as research reagents

CC for characterising sites of ligand-receptor interaction.

XX SQ Sequence 385 AA;
 Query Match 23.8%; Score 507.5; DB 20; Length 385;
 Best Local Similarity 31.4%; Pred. No. 1.3e-45;
 Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;
 QY 27 LLFLPVTVCOSGINVSDNSAKPTLTIK--SFNGGPONTFEEFPLSDIEGWTGATTIKAE 84
 DB 5 LLLWPLVL--GFSLSGGTQTPSVYDESGSTGGDDSTPSILPAP--RGYPGVQV----- 53
 QY 85 CPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYILLFVGVGPSNIVTLWKLSTKTSISLV 144
 DB 54 CANDS-DTLELPDSSRALLGWPTRLVPALYGLVVLVGLPANGALWLVLATQAPRLPST 112
 QY 145 IFHTNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVEMCRITTVFYGNNYCAILLITCMGI 204
 DB 113 MLMNLATADLLALALPPRIAYHLRGORWPFGEAACRLATAALYGHMYSVLLAAVSL 172
 QY 205 NRYLATAHPFTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVV 264
 DB 173 DRYLALVHPLRARALRGRLALGLCMAAWLMAAALALPLTLQRTQTLRLARSDRVLCHDAL 232
 QY 265 DACESPSSFRFYYVSVLAFFGLIPFVIIIFCYTTLIHKLKSKDRIMWGIKAVLLILVI 324
 DB 233 -PLDAQASHWQPAFTCLALLGCFPLLAMLLCYGATLHTLAASGRRYGHALRLTAVVLA 291
 QY 325 FTICFAPNTIILVIHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 DB 292 AVAFFVPSNLLLLHYSDPSPSANGNLYGAYVPSLALSTLNSCVDPFIYYVVS 344

RESULT 23

AAV45036
 ID AAV45036 standard; Protein; 385 AA.

XX AC AAV45036;

XX DT 31-MAY-2000 (first entry)

XX DE Human protease activated receptor-4.

XX KW Human; PAR-4; protease activated receptor;
 XX KW antisense molecule; PAR antibody; cytostatic; therapeutic;
 XX KW metastatic tumour cell; placental implantation; invasive cell.
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX FT Region 43..67

XX FT /note= "conserved unique region comprising protease
 XX FT activated domain and hirudin binding domain.
 XX FT Therapeutically useful antisense molecules are designed
 XX FT based on this region"

XX FT Region 228..230
 XX FT /note= "conserved residues in second extracellular
 XX FT loop used for designing therapeutically useful
 XX FT antisense molecules"

XX FN WO200008150-A1.

XX PD 17-FEB-2000.

XX PF 05-FEB-1999; 99WO-IL00079.

XX PR 07-AUG-1998; 98IL-0125698.

XX PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

XX PI Bar-Shavit R;

XX DR WPI; 2000-205706/18.

DR N-PSDB; AA250775.

XX Treating metastatic tumor cells useful for treating disorders involving
 XX placenta implantation in a female comprises administration of an
 XX antisense molecule complementary to an RNA sequence of a protease
 XX activated receptor protein -

PS Example 3; Fig 11b; 46pp; English.

XX The patent discloses a method to treat metastatic tumour cells using
 CC an antisense molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisense molecules and
 CC antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a human PAR-4
 CC protein used for producing antisense molecules for treating invasive
 CC cells.

XX SQ Sequence 385 AA;

Query Match 23.8%; Score 507.5; DB 21; Length 385;
 Best Local Similarity 31.4%; Pred. No. 1.3e-45;
 Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;

QY 27 LLFLPVTVCOSGINVSDNSAKPTLTIK--SFNGGPONTFEEFPLSDIEGWTGATTIKAE 84
 DB 5 LLLWPLVL--GFSLSGGTQTPSVYDESGSTGGDDSTPSILPAP--RGYPGVQV----- 53
 QY 85 CPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYILLFVGVGPSNIVTLWKLSTKTSISLV 144
 DB 54 CANDS-DTLELPDSSRALLGWPTRLVPALYGLVVLVGLPANGALWLVLATQAPRLPST 112
 QY 145 IFHTNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVEMCRITTVFYGNNYCAILLITCMGI 204
 DB 113 MLMNLATADLLALALPPRIAYHLRGORWPFGEAACRLATAALYGHMYSVLLAAVSL 172
 QY 205 NRYLATAHPFTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVV 264
 DB 173 DRYLALVHPLRARALRGRLALGLCMAAALALPLTLQRTQTLRLARSDRVLCHDAL 232
 QY 265 DACESPSSFRFYYVSVLAFFGLIPFVIIIFCYTTLIHKLKSKDRIMWGIKAVLLILVI 324
 DB 233 -PLDAQASHWQPAFTCLALLGCFPLLAMLLCYGATLHTLAASGRRYGHALRLTAVVLA 291
 QY 325 FTICFAPNTIILVIHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 DB 292 AVAFFVPSNLLLLHYSDPSPSANGNLYGAYVPSLALSTLNSCVDPFIYYVVS 344

RESULT 24

AAV47623
 ID AAV47623 standard; Protein; 385 AA.

XX AC AAV47623;

XX DT 21-JAN-2002 (first entry)

XX DE Human PAR4.

XX KW Human; protease-activated receptor; PAR4; N-terminal; activated PAR4;
 XX KW platelet aggregation; inhibition; tumour cell; proliferation.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX FT Peptide 48..53
 XX FT /label= N-terminal of activated PAR4

XX PN WO200158930-A1.

XX PD 16-AUG-2001.

XX PF 06-FEB-2001; 2001WO-US03807.

XX FH Key Location/Qualifiers
 FT Region 1..78
 FT Peptide /note= "Extracellular N-terminal region"
 FT Cleavage-site /note= "Signal peptide"
 FT Protein /note= "Cleaved by signal peptidase"
 FT Region 18..385
 FT Region /note= "Mature non-activated human PAR4"
 FT (AA50140) Modified-site 56
 FT Domain /note= "N-glycosylated"
 FT Region 79..102
 FT Region /note= "Transmembrane domain 1"
 FT Domain 103..110
 FT Region /note= "Intracellular loop 1"
 FT Domain 111..132
 FT Region /note= "Transmembrane domain 2"
 FT Region 133..150
 FT Domain /note= "Extracellular loop 1"
 FT Region 151..172
 FT Domain /note= "Transmembrane domain 3"
 FT Region 173..191
 FT Domain /note= "Intracellular loop 2"
 FT Region 192..213
 FT Domain /note= "Transmembrane domain 4"
 FT Region 214..239
 FT /note= "Extracellular loop 2, determines specificity for activating peptide"
 FT Region 228..230
 FT /note= "These three residues are conserved among PAR1-4"
 FT Domain 240..263
 FT Region /note= "Transmembrane domain 5"
 FT Domain 264..283
 FT Region /note= "Intracellular loop 3"
 FT Domain 284..305
 FT Region /note= "Transmembrane domain 6"
 FT Domain 306..316
 FT Region /note= "Extracellular loop 3"
 FT Region 317..343
 FT /note= "Transmembrane domain 7"
 FT Region 344..385
 FT /note= "Intracellular C-terminal region"
 XX WO9950415-A2.
 XX 07-OCT-1999.
 XX 31-MAR-1999; 99WO-US07100.
 XX 01-APR-1998; 98US-0053866.
 XX (ZYMO) ZYMOGENETICS INC.
 XX (UNIW) UNIV WASHINGTON.
 XX Xu W, Presnell SR, Yee DP, Foster DC;
 XX WPI; 1999-633640/54.
 XX Novel protease activated receptor 4, useful for screening for
 XX (ant)agonists for promoting the proliferation and/or differentiation of
 XX platelets and in mediating inflammatory events -
 XX Example 2; Page -; 85pp; English.
 XX This sequence represents a mutant human protease-activated receptor PAR4
 XX R47A, compared with PAR4 mutant R68A (AA50139) in studies to evaluate
 XX the importance of a protease cleavage site between residues 47 and 48.
 XX This mutant PAR4 was unable to be cleaved by thrombin or trypsin, in
 XX contrast to the wild-type (AA50135) or the R68A mutant. This cleavage
 XX activates PAR4, exposing a tethered hexapeptide ligand at the new

CC N-terminus. Protease-activated receptors (PARs) are a subfamily of G
 CC protein coupled receptors which are capable of mediating cellular
 CC signalling in response to proteases. Agonists of PAR4 are useful for
 CC upregulating cellular or physiological responses whereas antagonists
 CC are used to downregulate these activities. The PAR4 protein is
 CC further useful for dissecting the effects of thrombin or other
 CC activating proteases in the clotting pathway from the effects of these
 CC proteases at the cellular level. Agonists are specifically useful in
 CC promoting the proliferation and/or differentiation of platelets, in
 CC mediating inflammatory events, responses to vascular injury, chemotaxis
 CC or mitogenesis, and in producing growth factors. Antagonists are useful
 CC as research reagents for characterising sites of ligand-receptor
 CC interaction.
 CC Note: This sequence is not shown in the specification, but is derived
 CC from the sequence of the PAR4 shown in figure 1.
 XX SQ Sequence 385 AA;
 Query Match 23.7%; Score 506.5; DB 20; Length 385;
 Best Local Similarity 31.4%; Pred. No. 1.7e-45;
 Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;
 QY 27 LLPLPVTVCSQGINVSDNSAKPTLTIK--SFNGGPONTFEEPLSDIEGWTGATTIKAE 84
 DB 5 LLLWPLVL--GFSLSGGTQTSVYDESGSTGGDDSTPSILPAP--AGTPGV----- 53
 QY 85 CPEDSISTLHVNNATIGYLRSSLSITQVIPAIVILLFWGVPSNIVTLWKLSTKSLV 144
 DB 54 CANDS-DTLELPDSSRALLLGWVPTLVLPALYGLVGLPANGALWLVATQAPRLPST 112
 QY 145 IFHTNIAIADLLFCVTLTPFKIAYHLNGNWNVGEVNCRTITVVVFYNGMYCAIILTCMGI 204
 DB 113 MLMLNLATALLALPPRIAYHLRQRPFGAACLRLAALYGHMYGSVLLAAVSL 172
 QY 205 NRYLATAHPTTYQKLPKRSFLLMCGIVVMVFLYMLPFVILKQEHLYHSEITTCDDVV 264
 DB 173 DRYLALVHPLRALRGRRLALGLCNAWLMAALALPLTLQKQTFRLARSRVLCHDAL 232
 QY 265 DACESPSFRFYFVSLAFFGLIPFVILFVTTLIHKLSKDRWLGVIKAVLLILVI 324
 DB 233 -PLDAQASHWQPAFTCTALLGCFLLAMLLCYGATLHTLAASGRYGHRLTAVVLAS 291
 QY 325 FTICFAPTNIILVIHANYYYNHTSDLYFWYLIACLGSLNSCLDPFLYFVMS 377
 DB 292 AVAFFVPSNLLLLLHYSDPSAGNLYGAYVPSLALSTLNSCVDPFIYVVS 344
 RESULT 27
 ID AAW16314 standard; Protein; 892 AA.
 XX AAW16314;
 XX 16-AUG-1997 (first entry)
 XX Human thrombin receptor-Yeast G-alpha protein fusion.
 XX G-protein coupled receptor; agonist; antagonist; assay;
 XX G-alpha protein; Gpalp; GPAL gene; thrombin receptor.
 XX Chimaeric Homo sapiens;
 XX Chimaeric Saccharomyces cerevisiae.
 XX Key Location/Qualifiers
 XX Peptide 1..13
 XX /label= Sig_peptide
 XX /note= "alpha-factor (STE2) signal peptide"
 XX Protein 14..892
 XX /label= ThR-GPAL_fusion
 XX Region 14..417
 XX /label= ThR
 XX /note= "thrombin receptor amino acids 22-425"
 XX Peptide 418..420

FT Region /label= Linker
 FT 421..892
 FT /label= G-alpha_protein
 XX WO9711159-A1.
 XX 27-MAR-1997.
 XX 20-SEP-1996; 96WO-US15203.
 XX 20-SEP-1995; 95US-0004023.
 XX (HEAR-) HEARTLAND BIOTECHNOLOGIES LLC.
 XX Boulton TG, Das PO, Mandell RB, McMullen TW;
 XX WPI; 1997-202868/18.
 XX N-PSDB; AAT62461.
 XX New yeast constructs for use in screening assays - which express a
 XX fusion comprising a mammalian or fungal transmembrane receptor and
 XX a yeast G-alpha protein
 XX Example 2; Fig 12A-G; 109pp; English.
 XX A fusion protein (AAW16314) comprises the human thrombin receptor
 XX (amino acids 22 to C-terminal threonine) covalently linked to
 XX yeast G-alpha protein spaip (encoded by the GPA1 gene). It is
 XX expressed by a thr-GPA1 gene fusion (AAT62461) in vector PRMH816.
 XX Thrombin-dependent activation of the yeast mating pathway by the
 XX Thr-GPA1 fusion protein was observed in ste2 gpal- yeast cells.
 XX The polypeptide is an example of novel fusion proteins between
 XX eukaryotic heterotrimeric G-protein coupled receptors and yeast
 XX G-alpha proteins that can be coupled to the pheromone-induced
 XX signal transduction pathway of yeast and used in screening assays
 XX to identify agonists or antagonists of the receptor.
 XX
 SQ Sequence 892 AA;
 Query Match 23.5%; Score 502; DB 18; Length 892;
 Best Local Similarity 32.4%; Pred. No. 1.7e-44;
 Matches 114; Conservative 75; Mismatches 127; Indels 36; Gaps 12;
 Qy 49 TLTKSF-NGGPQNTFFEFPLSDIEGWTGAT----TTIKACPEDSISTLHVNNATIGYL 103
 Db 29 TLDPRSELLRNPNDKYEPFWEDEKESGLTEYRLVSINKSPLOKQLPAFISDASGYL 88
 Qy 104 RSSLSQVIPAIIYILLFVVGVPNSI--VTLWKLRLTKSISLVIFHTNLAIADLLFCVTL 161
 Db 89 TSSWLTFLVPSVYTGCVFVSLPLNMAIVVFIKMKVKK-PAVVYMLHLATADVLFVSVL 147
 Qy 162 PFKIAYHLGNWVGEVWMCRIITVVFYGNMYCAIILTCGINRYLATAHP---FTYQK 218
 Db 148 PFKISYFSGDWQFGSELRCFVFAAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRT 207
 Qy 219 LPKSFSLMCGIYVWVFLVYMLPFVILKQBYHLVHSEITTCDDVDACSPSPSFREYF 278
 Db 208 LGRASFT---CLAWALAIAGVFLVLEKEQIQVPGNLITTCDDVNETLEGYYAYFS 264
 Qy 279 VSLAFFGFLIPFVLIICYTLIHLKS-----KDWIWLGYIKAVLL---ILVIFTI 327
 Db 265 AFSAVF-FFVPLIISTVCYVGIICLSSSAVANRKKSR-----ALFLSAAVFCIFII 316
 Qy 328 CFAPTNILVTHANYVH--NTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 Db 317 CFGPTNVLIIAHY-SFLSHTTSTAAPAYLLCVSSISSCIDPLIYYAS 367
 RESULT 28
 ABG35300
 ID ABG35300 standard; Protein; 425 AA.
 XX
 AC ABG35300;

XX 15-JUL-2002 (first entry)
 XX Human PAR1 type thrombin receptor.
 XX Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
 XX inflammatory disease; cell proliferative disease.
 XX Homo sapiens.
 XX JP2002010784-A.
 XX 15-JAN-2002.
 XX 29-JUN-2000; 2000JP-0196514.
 XX 29-JUN-2000; 2000JP-0196514.
 XX (TEIJ) TEIJIN LTD.
 XX WPI; 2002-321520/36.
 XX N-PSDB; ABK70889.
 XX An inhibitor of cell growth mediated by thrombin used to treat
 XX inflammatory and cell proliferative diseases -
 XX Disclosure; Page 26-27; 44pp; Japanese.
 XX The invention relates to a polypeptide or a compound which can inhibit
 XX cell growth caused by thrombin. The polypeptide/compound combines to a
 XX specific region of the structure of PAR1 type human thrombin receptor
 XX participating to cell growth. Preferably, the compound contains the
 XX 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
 XX human thrombin receptor ((X4)-Iyr-Glu-Pro-Phe-Trp-(X5) X4; X5 = optional
 XX amino acid or peptide sequence). Also included are a modified PAR1 type
 XX thrombin receptor gene or its fragment used for obtaining the above
 XX polypeptide, a human PAR1 type thrombin receptor protein and its
 XX encoding DNA comprising a fully. The polypeptide or the compound is used
 XX to treat inflammatory diseases and cell proliferative diseases. The
 XX present sequence is a PAR1 type thrombin receptor (or a modified
 XX version).
 XX
 SQ Sequence 425 AA;
 Query Match 23.5%; Score 501; DB 23; Length 425;
 Best Local Similarity 30.7%; Pred. No. 7.6e-45;
 Matches 119; Conservative 83; Mismatches 138; Indels 48; Gaps 14;
 Qy 20 LILVAAGLLPLPVTVCQSGINVSNDNAKP-----TLTKSF-NGGPQNTFFEFPLSDI 71
 Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRPSLLRNPNDKYEPFWEDEE 60
 Qy 72 EGWTGAT----TTIKACPEDSISTLHVNNATIGYRLSSLSQTVIPAIIYILLFVVGVPNSI 127
 Db 61 KNESGLTEYRLVSINKSPLOKQLPAFISDASGYLTSSWLTFLVPSVYTGCVFVSLPLN 120
 Qy 128 I--VTLWKLRLTKSISLVIFHTNLAIADLLFCVTLPFKIAYHLGNWVGEVWMCRIIT 185
 Db 121 IMAIVVFIKMKVKK-PAVVYMLHLATADVLFVSVLPFKISYFSGDWQFGSELRCFVT 179
 Qy 186 VVFYGNMYCAIILTCGINRYLATAHP---FTYQKLPKSFSLMCGIYVWVFLYMLP 242
 Db 180 AAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRTLGRASFT---CLAWALAIAGVVP 236
 Qy 243 FVILKQBYHLVHSEITTCDDVDACSPSPSFREYFVSLAFFGFLIPFVLIICYTYLIIH 302
 Db 237 LLLKEQTIQVPGNLITTCDDVNETLEGYYAYFSAFSAVF-FFVPLIISTVCYVGIIR 295
 Qy 303 KLKS-----KDWIWLGYIKAVLL---ILVIFTICFAPTNILVTHANYVH--NTD 349
 Db 296 CLSSSAVANRKKSR-----ALFLSAAVFCIFIIICFGPTNVLIIAHY-SFLSHTTTE 347
 Qy 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377

QY 348 TDSLYFWYLYALCLGSLNSCLDPLFYFVMS 377
Db 327 WGNLYGAYVPSLSTLNSCLDPLFYFVMS 356

RESULT 32
AAY84815
ID AAY84815 standard; Protein; 359 AA.
XX AC AAY84815;
XX DT 08-AUG-2000 (first entry)
XX DE A human G-protein coupled receptor designated HG52.
XX KW Human; G-protein coupled receptor; HG52; immune system; thrombin;
KW fibrinogen; fibrin; clotting factor; procoagulant; platelet activation;
KW chemotaxis; mitogenesis.
XX OS Homo sapiens.
XX PN WO200020438-A1.
XX PD 13-APR-2000.
XX PF 29-SEP-1999; 99WO-US22634.
XX PR 02-OCT-1998; 98US-0102958.
XX PA (MERI) MERCK & CO INC.
XX PI Liu Q, McDonald TP, Wang R;
XX DR WPI; 2000-317696/27.
XX N-PSDB; AAA14828.

New recombinant DNA encoding a G-protein coupled receptor designated
HG52 is useful to find modulators of thrombin effects and shows
homology to thrombin receptors -
Claim 6; Fig 2; 39pp; English.

The present sequence represents a human G-protein coupled receptor
designated HG52. HG52 is a member of the rhodopsin family. The HG52
RNA is widely expressed in humans as a transcript of about 4.5 kb,
especially in cells of the immune system. The HG52 DNA can be used
in chromosomal mapping studies, and to identify individuals carrying
a disease-carrying gene. Agonists and antagonists of HG52 will be
useful as modulators of the effects of thrombin, including conversion
of fibrinogen to fibrin in plasma, activation of clotting factors V,
VIII, XIII and protein C, the procoagulant function of platelets and
endothelial cells, stimulation of platelet activation, chemotaxis of
monocytes and lymphocytes, and mitogenesis of lymphocytes and
mesenchymal cells such as vascular smooth muscle cells, fibroblasts
and epithelial cells.

Sequence 359 AA;
Query Match 23.1%; Score 493; DB 21; Length 359;
Best Local Similarity 35.7%; Pred. No. 4.3e-44;
Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;

QY 96 NNATIGYLRSLSTQVPIAYILLFVGVPSNIVTLWKLSTKXIS-LVIFHTNLAIAD 154
Db 10 DNATLQMLRNPAIALPVVLSVAASIPGNLSLWLCRMGPRSPSVIFMNLSTVD 69
QY 155 LLFCVTLFPKAYHLNGNNWVFGVMCRITTVFVGNMYCAILITCMGIRYLATARPF 214
Db 70 LMLASVLPFQIYYHCNRHHWVGVLLCNVTVFAVANNYSILMTWTCISVERFLGVLYPL 129
QY 215 TYQKLPKRSFSLMCGIWMVWVFLVMLPFLVKQYHLVHSEITTCDDVDACSPS--- 271
Db 130 SSKRRRRRYAAACAGTWMLLLLTALSPLARTDITVPVHALGIITCFDVLKWTMLPSVAM 189

QY 272 --SFREYFVSLAFFGFLIPFVIIIFCYTTLIHKL-----KSDRIWLGYIKAVLLI 321
Db 190 WAVFLFTIFILL----FLIPFVITVACYTATILKLRTEEAHGREQRRRAVGLAAVVLL- 244
QY 322 LVIFTTCFAPTNILVIHHANYHYNTDSLYFMYLALCLGSLNSCLDPLFYFVMSK 378
Db 245 --AFVTCFAPNPFVLLAHIVLSRLFYG-KSYHYVYKLTCLCLNCLNCLDPLFYFVMSR 298

RESULT 33
AAG78530
ID AAG78530 standard; Protein; 359 AA.
XX AC AAG78530;
XX DT 12-MAR-2002 (first entry)
XX DE G-protein coupled receptor (designated Paul) amino acid sequence.
XX KW G-protein coupled receptor; Paul; 7TM receptor; GPC receptor;
KW human; antibacterial; virucide; fungicide; protozoacide;
KW anti-Human Immunodeficiency Virus; analgesic; cytostatic; antidiabetic;
KW anorectic; antiasthmatic; antiparkinsonian; cardiant; hypertensive;
KW hypotensive; osteopathic; antianginal; antimanic; cerebroprotective;
KW antitumor; antiallergic; antidepressant; antimigraine; antiemetic;
KW tranquiliser; neuroleptic; neuroprotective; nootropic; anticonvulsant;
KW Human Immunodeficiency virus type 1; HIV-1; HIV-2; pain; cancer;
KW diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease;
KW acute heart failure; hypotension; hypertension; urinary retention;
KW osteoporosis; angina pectoris; myocardial infarction; stroke; ulcer;
KW allergy; benign prostatic hypertrophy; migraine; vomiting; anxiety;
KW schizophrenia; manic depression; delirium; dementia; mental retardation;
KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.
XX OS Homo sapiens.
XX PN US2001029032-A1.
XX PD 11-OCT-2001.
XX PF 18-DEC-2000; 2000US-0739151.
XX PR 09-OCT-1998; 98US-103789P.
XX PR 06-OCT-1999; 99US-0413534.
XX PA (ZHU)/ ZHU Y.
XX PA (LIXX)/ LI X.
XX PA (VAVT)/ VAWTER L.
XX PI Zhu Y, Li X, Vawter L;
XX DR WPI; 2001-647985/74.
XX N-PSDB; AA164231.

New G-protein coupled receptor polypeptide, referred as Paul and
encoding polynucleotide, useful for diagnosing and treating cancers,
infections, neurological disorders, diabetes, asthma and identifying
modulators -
Claim 1c; Page 13-14; 15pp; English.

The invention relates to an isolated polypeptide, a member of G-protein
coupled receptor family of polypeptides, comprising a fully defined
sequence of 359 amino acids, its 95% identical sequence, or a polypeptide
encoded by a polynucleotide comprising a fully defined sequence of 1080
base pairs defined in the specification, or a fragment or variant of it.
The activity of the protein of the invention may be described as
antibacterial, virucide, fungicide, protozoacide, anti-Human
Immunodeficiency virus, analgesic, cytostatic, antidiabetic, anorectic,
antiasthmatic, antiparkinsonian, cardiant, hypertensive, hypotensive,
osteopathic, antianginal, antimanic, cerebroprotective, antitumor,
antiallergic, antidepressant, antimigraine, antiemetic, tranquiliser,

neuroleptic, neuroprotective, nootropic and anticonvulsant. Polypeptides and polynucleotides of the invention are useful in diagnosis and in identifying compounds such as agonists and antagonists which are useful in therapy. They are also useful for treating diseases, including infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by Human Immunodeficiency virus type 1 or 2 (HIV-1 or HIV-2), pain, cancers, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypertension, hypotension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers, allergies, benign prostatic hypertrophy, migraine, vomiting, psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, delirium, dementia, and severe mental retardation, dyskinesias such as Huntington's disease or Gilles de la Tourette's syndrome. Polypeptides and polynucleotides of the invention may also be useful as diagnostic reagents, for example in the detection of mutations in the associated gene, chromosome localisation studies and expression pattern determination. The current sequence represents a G-protein coupled receptor (designated "Paul" in the specification) amino acid sequence.

Sequence 359 AA:
SQ

Query Match	23.1%	Score	493;	DB	22;	Length	359;
Best Local Similarity	35.7%;	Pred.	No. 4.3e-44;				
Matches	106;	Conservative	53;	Mismatches	116;	Indels	22;
						Gaps	6
<hr/>							
Qy	96	NNATGYLRSSLS	STQVTPAIYILLFVVGVS	PNVTLWKL	SLR	TKSTS	-LVIFHTNLATAD 154
		:	:	:	:	:	:
Db	10	DNATLQMLRNP	AIAPVVSLS	AAVSIPGNL	FSLWVL	CRRRGPRSP	SVIEMINLSVTD 69
		:	:	:	:	:	:
Qy	155	LLFCVTLPPK	YAYHLGNMNV	GEV	MCRIITVV	FYGNMYCAI	LILTCMGINYLATAHPF 214
		:	:	:	:	:	:
Db	70	LMASVLFPQ	IYHCNRHNV	GGVLLCN	VNVTVAFY	ANMYSSILIT	MTCISVERFLGVLVPL 129
		:	:	:	:	:	:
Qy	215	TYQKLPKRS	FSLLMCGIY	VWVFLY	MLPFLV	LKQEVHLV	HSITTCDDVVVDACEPS--- 271
		:	:	:	:	:	:
Db	130	SSKRWRR	RYAVACAGT	WLLLLT	ALSP	LARTDLY	FPVHALGITCFDVLKWTMLPFSVAM 189
		:	:	:	:	:	:
Qy	272	--SFRFYV	SLAPFGFL	PIPVII	IFCYTTL	LIHKL	-----KSKDRILWLGIVKAVILLI 321
		:	:	:	:	:	:
Db	190	WAVLFTT	FILL---	FLIP	PVIT	VACYTATIL	KLRTTEBAHQREQRRAVGLAAVLL- 244
		:	:	:	:	:	:
Qy	322	LVIFTCF	APNWI	LVIH	HHYNYH	NTDSLY	FWMYLIALCLGSLNUSCLDPLFYVMSK 378
		:	:	:	:	:	:
Db	245	--AFVTC	APNFWLL	AHIV	SRLEG-	KSYHVYKLT	CLCLSNLNCDFVYFYFASR 298
		:	:	:	:	:	:

```

PF 16-NOV-2000; 2000WO-US31581.
XX
XX 16-NOV-1999; 99US-0165838.
XX 17-NOV-1999; 99US-0166071.
XX 19-NOV-1999; 99US-0166678.
XX 28-DEC-1999; 99US-0173396.
XX 22-FEB-2000; 2000US-0184129.
XX 28-FEB-2000; 2000US-0185421.
XX 28-FEB-2000; 2000US-0185421.
XX 02-MAR-2000; 2000US-0186530.
XX 03-MAR-2000; 2000US-0186811.
XX 09-MAR-2000; 2000US-0188114.
XX 17-MAR-2000; 2000US-0190310.
XX 21-MAR-2000; 2000US-0190800.
XX 20-APR-2000; 2000US-0198568.
XX 02-MAY-2000; 2000US-0201190.
XX 08-MAY-2000; 2000US-0203111.
XX 25-MAY-2000; 2000US-0207094.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejltz T, Huff RM;
XX
XX WPI; 2001-389826/41.
XX N-PSDB; AAH51006.
XX
XX New G protein-coupled receptor (ngPCR-x) and its encoding
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX
XX Claim 37; Page 88; 261pp; English.
XX
XX The present invention relates to novel G protein-coupled receptors
CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC sequence is one such G protein-coupled receptor. GPCRs are also known as
CC seven transmembrane receptors and function in signal transduction. The
CC ngPCRx coding sequences are useful for screening a human to diagnose a
CC disorder affecting the brain or a genetic predisposition, specifically
CC schizophrenia. ngPCRx are useful for identifying compounds useful for
CC treating schizophrenia. Detection of ngPCRx in a sample is useful as a
CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
CC metabolic and cardiovascular diseases, proliferative disorders and
CC hormonal disorders. Modulators of ngPCRx activity have the utility for
CC treating neurological disorders, including schizophrenia, ADHD/ADD
CC (attention deficit-hyperactivity disorder/attention deficit disorder),
CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
CC migraine and senile dementia. Additional disorders include inflammatory
CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
CC diseases e.g. inflammatory bowel disease.
XX
XX Sequence 359 AA;
XX
XX Query Match 23.1%; Score 493; DB 22; Length 359;
XX Best Local Similarity 35.7%; Pred. No. 4.3e-44;
XX Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6
XX
Qy 96 NNATIGYRRSLSTQVIPAIIYLLFFVGVGPSNIVTLWKLIRTKSIS-LVIFHTNLAIA 154
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
10 DNATLQMLRNPATAVALPVVYSLVAAVSIFGNLFSVLVLCRRMGPRSPSVIFMNLVSVD 69
:|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Qy 155 LLFCVTLPFKIAIHLNNGNNVFEVMCRITTVFVGYGMYCAILLTCWGINRYIATAHPF 214
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
70 LMLASVLFPFOIYHCNRHHWVFGVLLCNVVTAFYANNYSILMTCTISVERFLGVLYPL 129
:|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Qy 215 TYOKLKRPSLMLCGIWWVVFVLYMLPFLVILKQYHLVHSEITTCDDVVDACESPS-- 271
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
130 SSRWRRRRYAVACACTWLLLTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189
:|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Qy 272 --SFRFYFVSLAFFGFLIFPFIIFCYTTLIHKL-----KSKDRIMWLGYIKAVLLI 321
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

```

Db 190 WAVFLFTIFILL----FLIPFVITVAVTATILKLRTEAHGRRRAVGLAAVVLL- 244

Qy 322 LVIFTICFAPNTIILVIHANYHYNTDSLYFMVLIALLCLGSLNSCLDPELYFVMSK 378

Db 245 --AFVTCFAPNFFVLLAHIVSRFLYG-KSYHYVYKLTLCLSCLNCLDPEVYFASR 298

RESULT 35
AAB62285
ID AAB62285 standard; Protein; 359 AA.
XX
AC AAB62285;
XX
DT 29-JUN-2001 (first entry)
DE Human G-protein coupled receptor, PAUL.
XX
KW G-coupled protein receptor; PAUL; anti-HIV, antibacterial; antiviral;
KW antifungal; protozoacide; cytostatic; antidiabetic; anorectic; human;
KW antiaesthetic; antiparkinsonian; cardiac; hypertensive; osteopathic;
KW antianimal; cerebroprotective; antitumor; antimigraine; antineurotic;
KW tranquilizer; nootropic; anticonvulsant; neuroleptic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200125280-A1.
XX
PD 12-APR-2001.
XX
PF 03-OCT-2000; 2000WO-US27228.
XX
PR 06-OCT-1999; 99US-0413534.
XX
PA (SMIK) SMITHLINE BEECHAM CORP.
XX
PI Zhu Y, Li X, Yawter L;
XX
DR WPI; 2001-273570/28.
DR N-PSDB; AAF57649.
XX
XX Novel G-coupled protein receptor, PAUL useful for treating diseases
PT such as microbial infections, cancers, obesity, asthma, diabetes,
PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer and
PT allergy -
XX
PS Claim 1; Page 25; 34pp; English.
XX
CC This represents a human G-coupled protein receptor (designated PAUL).
CC The PAUL polypeptide and polynucleotide are useful for treating and
CC diagnosing infections such as bacterial, fungal, protozoan and viral
CC infections, particularly infections caused by human immunodeficiency
CC virus (HIV)-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia,
CC bulimia, asthma, Parkinson's disease, acute heart failure, hypotension,
CC hypertension, urinary retention, osteoporosis, angina pectoris, stroke,
CC myocardial infarction, ulcers, allergies, benign prostatic hypertrophy,
CC migraine, vomiting, psychotic and neurological disorders including
CC anxiety, schizophrenia, manic depression, depression, delirium, dementia
CC and severe mental retardation, and dyskinesia such as Huntington's
CC disease or Gilles de la Tourette's syndrome. The PAUL sequences are
CC useful as vaccines to induce an immunological response in a mammal. The
CC PAUL polypeptide is also useful for identifying its modulators. The PAUL
CC polynucleotide is valuable for chromosome localization studies and for
CC tissue expression studies.
XX
SQ Sequence 359 AA;
Query Match 23.1%; Score 493; DB 22; Length 359;
Best Local Similarity 35.7%; Pred. NO. 4.3e-44;
Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;

Qy 96 NNATIGYRSLSTQVIPAIVLLFVGVPSNIVTLKLSLRTKXIS-LVIFHTNLAID 154

Db 10 DNATQLMRNPAIAVLEPVVYSLVAASVIGNLSFVLCRRMGPRSPSIFMNLVSVD 69

Qy 155 LLFCVTLPPFKIAYHLNGNNVFGVEMCRITTVFYGNMYCAILILTCMGINRYLATAHPF 214

Db 70 LMLASVLPQIYVHCNRRHFWGVLLCNVTVVAFYANMYSSILMTCTISVERELGVLVPL 129

Qy 215 TYOKLPKRSFSLMCGIIVWVFLYMLPFVILKQEHVHSEITTCDDVVDACESPS--- 271

Db 130 SSKRWRRRYAAACAGTLLLLTALSPLARTDLYPVHALGIITCFDVLKWTMLPSVAM 189

Qy 272 --SFREYFVSLAFAFGFLIPFVIIIFCYTTLIHL-----KSDRIWLGVYKAVLLI 321

Db 190 WAVFLFTIFILL----FLIPFVITVAVTATILKLRTEAHGRRRAVGLAAVVLL- 244

Qy 322 LVIFTICFAPNTIILVIHANYHYNTDSLYFMVLIALLCLGSLNSCLDPELYFVMSK 378

Db 245 --AFVTCFAPNFFVLLAHIVSRFLYG-KSYHYVYKLTLCLSCLNCLDPEVYFASR 298

RESULT 36
AAY69485
ID AAY69485 standard; Protein; 359 AA.
XX
AC AAY69485;
XX
DT 03-JUL-2000 (first entry)
DE Amino acid sequence of a human 14400 receptor polypeptide.
XX
KW Human; 14400 receptor; G-protein coupled receptor; signalling pathway;
KW CD8 T cell; CD34+ bone marrow cell; thrombocytopenia;
KW inflammation; spleen disorder; splenomegaly; lung disorder;
KW adult respiratory distress syndrome; colon disorder;
KW bacterial enterocolitis; liver disorder; hepatic injury;
KW platelet number; precursor T-cell neoplasm.
XX
OS Homo sapiens.
XX
FH Key
FT Domain 1..23
FT Domain /note= "extracellular domain"
FT Domain 24..256
FT Domain /note= "transmembrane domain"
FT Domain 120..122
FT Misc-difference 278 /note= "GPCR signal transduction signature"
FT FT /note= "Thr encoded by AGC"
FT Domain 297..359
FT Domain /note= "intracellular domain"
XX
WO200011170-A1.
XX
02-MAR-2000.
XX
20-AUG-1999; 99WO-US19112.
XX
20-AUG-1998; 98US-0137063.
XX
20-AUG-1999; 99US-0378100.
XX
(MILL-) MILLENNIUM PHARM INC.
XX
Glucksmann MA, Weich NS;
XX
WPI; 2000-256378/22.
DR N-PSDB; AAZ99588.
XX
G-protein coupled receptor, 14400 receptor, for treating CD4, CD8 or
CD34+ related disorders -
XX
Claim 1; Fig 1A-B; 92pp; English.
XX
The present sequence represents a human 14400 receptor polypeptide. The
receptor is a G-protein coupled receptor that participates in signalling
pathways. The polypeptide can be used to identify an agent to modulate

CC the polypeptides activity in patients having a disorder involving CD8 or
CC CD4 T cells or CD34+ bone marrow cells, especially where the disorder
CC involves thrombocytopenia or inflammation. The modulators can
CC be used to treat disorders involving the spleen (e.g. splenomegaly),
CC lung (e.g. adult respiratory distress syndrome), colon (e.g. bacterial
CC enterocolitis), liver (e.g. hepatic injury), the uterus and
CC endometrium, brain, T-cells, skin, heart, blood vessels, red cells,
CC thymus, B-cells, kidney, breast, testis and epididymis, prostate,
CC thyroid, skeletal muscle, pancreas, small intestine, reduced platelet
CC number, precursor T-cell neoplasms, CD3, CD4, and CD8 T lymphocytes.
XX
SQ Sequence 359 AA;

Query Match 22.9%; Score 489; DB 21; Length 359;
Best Local Similarity 34.7%; Pred. No. 1.2e-43;
Matches 102; Conservative 55; Mismatches 121; Indels 16; Gaps 5;
QY 96 NNATIGYRSLSTQVPAIYILLFVVGVPNSIVTLWKLSTRTKSIS-LVIFHTNLAIAD 154
DB 10 DNATQLMRNPAIAVALPVVYSLVAASIPGNLFLSLWLCRRMGPRSPSVIFMNLSTVD 69
QY 155 LLFCVTLPFKTAYHLNGNNVFGVEMCRITTVVPGYGNMYCAILLTCGNNRYLATAPF 214
DB 70 LMLASVLFPQIYXHCNRRHWHVGLLCNVVTVAFYANNYSILMTCTISVERFLGVLYPL 129
QY 215 TYQKLPKRSPLLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVVDACESPS- 271
DB 130 SSKWRRRRYAAVACAGTWLILLTALSPLARTDLYPHALGIITCFDVLKWTMLPSVAM 189
QY 272 --SRFYFVSLAFGLFPFVITIFCYTTLIHLKSKDRIW-----LGVIKAVLLILVI 324
DB 190 WAVEFLTIFILL-----FLIPPVITVACVATILKLRTEEAHGREQSAAGVLLA 245
QY 325 FTICFAPTNILVTHANNYYHNTDSLFWMLIALCLGSLNSCLDPFLYFVMSK 378
DB 246 FVTCFAPNFFLLAHIVSRLEFYG-KSYHYVYKTLCLCLNCLDPFVYFASR 298

RESULT 37
ABG35298
ID ABG35298 standard; Protein; 402 AA.
XX
AC ABG35298;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PAR1 type thrombin receptor delta 1-49TR.
XX
KW Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX
OS Homo sapiens.
XX
FN JP2002010784-A.
XX
PD 15-JAN-2002.
XX
PF 29-JUN-2000; 2000JP-0196514.
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
PA (TEIJ) TEIJIN LTD.
XX
DR WPI; 2002-321520/36.
DR N-PSDB; ABK70887.
XX
PT An inhibitor of cell growth mediated by thrombin used to treat
XX inflammatory and cell proliferative diseases -
PS Disclosure; Page 22-24; 44pp; Japanese.
XX
CC The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a

CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Iyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
CC version).
XX
SQ Sequence 402 AA;

Query Match 22.7%; Score 485.5; DB 23; Length 402;
Best Local Similarity 31.8%; Pred. No. 3.2e-43;
Matches 107; Conservative 74; Mismatches 121; Indels 35; Gaps 11;
QY 63 FEEFPLSDIEGWTGAT----TTIKACPEDSISTLHVNNTATIGYLRSSLSLTQVIPAIIYL 118
DB 29 YEFWEDEEKNESEGLTEYRLVINSKSPLOKQLPAFISEDASGYLTSSWTLTFVPSVYTG 88
QY 119 LFVGVGPSNI--VTLWKLSTRTKSISLVIFHTNLAIADLLFCVTLFPKIAHYHLNGNNWVF 176
DB 89 FVVVSLPLNIMATVWVFLKMKVKK-PAVVVMLHLATADLVFVSVLPFKISYFSGSDWQF 147
QY 177 GEVWCRTITVFGNNMYCAILLTCGNNRYLATAPF---FTYQKLPKRSFSLLMCGIWM 233
DB 148 GSELCHFVTAAPYCNMYASILLMTVISIDRFLLAVVYPMOSLSWRTIGRASFT---CLAIW 204
QY 234 VMVFLMPLPFVILKQYHLVHSEITTCDDVVDACESPSFRFYFVSLAFGLFPFVII 293
DB 205 ALAIAGVVPLLLKEQIIVQVGLNITTCDDVNLTELEGYAYVYFSAFV-FVPLIIS 263
QY 294 IFCYTTLIHLKKS-----KDRIWGIKAVLL---ILVITFCFAPTNILVIHAN 342
DB 264 TVCVSVSIIRCLSSSAVANRSKSR-----ALFLSAAVFCIFICGPTNVLIIAHY-S 315
QY 343 YYVH--NTDSLFWMLIALCLGSLNSCLDPFLYFVMS 377
DB 316 FLSHTSTTEAAYFAYILLVCVSVSISCCIDPLIIVYAS 352

RESULT 38
ABG35299
ID ABG35299 standard; Protein; 371 AA.
XX
AC ABG35299;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PAR1 type thrombin receptor delta 1-80.
XX
KW Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX
OS Homo sapiens.
XX
FN JP2002010784-A.
XX
PD 15-JAN-2002.
XX
PF 29-JUN-2000; 2000JP-0196514.
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
PA (TEIJ) TEIJIN LTD.
XX
DR WPI; 2002-321520/36.
DR N-PSDB; ABK70888.
XX
PT An inhibitor of cell growth mediated by thrombin used to treat
XX inflammatory and cell proliferative diseases -

XX Disclosure; Page 24-25; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit

CC cell growth caused by thrombin. The polypeptide/compound combines to a

CC specific region of the structure of PAR1 type human thrombin receptor

CC participating to cell growth. Preferably, the compound contains the

CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type

CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional

CC amino acid or peptide sequence). Also included are a modified PAR1 type

CC thrombin receptor gene or its fragment used for obtaining the above

CC polypeptide, a human PAR1 type thrombin receptor protein and its

CC encoding DNA comprising a fully. The polypeptide or the compound is used

CC to treat inflammatory diseases and cell proliferative diseases. The

CC present sequence is a PAR1 type thrombin receptor (or a modified

CC version).

XX Sequence 371 AA;

Query Match 22.7%; Score 484.5; DB 23; Length 371;

Best Local Similarity 34.2%; Pred. No. 3.7e-43;

Matches 101; Conservative 67; Mismatches 96; Indels 31; Gaps 10;

QY 101 GYLSSLSSTQVPAIYILLFVGVPSNI--VTLMKLSLRKTSISLVIFHTNLATADLLFC 158

DB 40 GYLTSWLTUFVPSVYTGVSPLTMAIWFILKMKVKK-PAVVMYMLLATADVLV 98

QY 159 VTLFPFKIAYHLNGNNWVFGVMCRITTVFVGNMYCAILITCMGINRYLATAPHP---FT 215

DB 99 SVLPFKISYFSGSDWQFSGELCFVTAFAFCNNYASILLMTVISIDRFVAVYPMQSL 158

QY 216 YQKLPKRSFSLLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVDACESPSPSRF 275

DB 159 WRTLGASFT---CLAIWALAIAGWPLLLKEQTIQVPLNITTCDDVNLLEGGYAY 215

QY 276 YFVSLAFPGFLIPFVIIICFTYTLIHKLKS-----KDRIWGLYKAVLL---TLVI 324

DB 216 YFSAFSAVF-FFVPLIISTVCYVSIIRCLSSAVANRSKSR-----ALFLSAAVFCI 267

QY 325 FTICFAPTNIILVHHANYHH--NTDSLYFMVLIALLCLGSLNSCLDPFLYFVMS 377

DB 268 FIICFGTNNVLIHAY-SFLSHTSTTEAAYFAYLLVCVSSISCCIDPLIYYAS 321

RESULT 39

AAW69598

ID AAW69598 standard; Protein; 359 AA.

XX

AC AAW69598;

XX

16-OCT-1998 (first entry)

DT

DE Mouse G-protein coupled receptor 69A08 #2 protein.

XX

XX Mouse; CKDLR201.1; chemokine; G-protein coupled receptor; GPCR;

KW 7 transmembrane receptor; inflammation; asthma; antiviral;

KW abnormal cell proliferation; regeneration; degeneration; atrophy.

XX

OS Mus sp.

XX

PN WO9831810-A2.

XX

23-JUL-1998.

XX

20-JAN-1998; 98WO-US00218.

XX

21-JAN-1997; 97US-0786624.

XX

PA (SCHE) SCHERING CORP.

XX

PI Capone M, Gorman DM, Hedrick JA, Huffine CF, Rossi DL;

XX Vicari A, Zlotnik A;

XX

DR WPI; 1998-414108/35.

XX N-PSDB; AAV40373.

PT Pure or recombinant chemokine CKDLR201.1 - useful, e.g. for treating

XX inflammation and as antiviral agents

PS Claim 2; Page 68-69; 77pp; English.

XX

CC The present sequence is mouse G-protein coupled receptor 69A08 #2

CC protein which comprises a plurality of epitopes. Host cells containing to

CC vectors comprising a nucleotide sequence encoding the protein are used to

CC produce recombinant protein. Treatment of a cell, particularly neuron,

CC macrophage or lymphocyte, carrying a G-protein coupled receptor (GPCR),

CC or a receptor responsive to CKDLR201.1 chemokine protein, with an

CC (ant)agonist is used to control physiological development, e.g.

CC alteration of calcium ion influx, a chemoattractant response, morphology,

CC phosphoinositide lipid turnover or an antiviral response. Nucleotide

CC sequences the CKDLR201.1 protein, or GPCRs, are useful as primers or

CC probes, e.g. for detecting and isolating related sequences and for

CC expressing antigenic peptides. Antibodies (Ab) directed against the

CC CKDLR201.1 protein and GPCRs are used to detect or purify the proteins;

CC diagnostically (e.g. for developmental abnormalities); in screening for

CC potential drugs; to inhibit chemokine/receptor activation; (when coupled

CC to a toxin or radioisotope) for killing specific cells, and to raise

CC anti-idiotype antibodies. CKDLR201.1 protein and GPCRs and compounds

CC which bind them can be used to treat inflammation, e.g. asthma; as

CC antiviral agents, and to treat abnormal cell proliferation, regeneration,

CC degeneration and atrophy. Therapeutic agents are administered orally, by

CC injection and rectally.

XX Sequence 359 AA;

Query Match 20.8%; Score 444.5; DB 19; Length 359;

Best Local Similarity 33.3%; Pred. No. 7e-39;

Matches 93; Conservative 57; Mismatches 110; Indels 19; Gaps 4;

QY 108 STQVIPAIIYILLFVGVPSNIIVTLWKLSLRKTSISLVIFHTNLAIADLLFCVTLPKIAY 167

DB 51 SPQLVPALYGLVAVGLPANGALWLVLRPLPSTILLMLNLAVADLLALVLPRLAY 110

QY 168 HLGNNWVFGVMCRITTVFVGNMYCAILITCMGINRYLATAPHPYKLPKRSFSL 227

DB 111 HLRGQWPFGAEACRVATAALYGHMVGSLVLLAAVSDRLVLPURARALRGQRUTTG 170

QY 228 MCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVDAC-----ESPSSRFYF 278

DB 171 LCLVAVLSAATLALPLTLHRQNFLL-APIACC--VMMRCPWLSRTPGTGERSAWLSWA 227

QY 279 VSLAFPGFLIPFVIIICFTYTLIHKLKSKDRIWGLYKAVLLILVITTCFAPTNIILVI 338

DB 228 ASL-----PLLAMGLCYGTTTTERALAANGQRYSHALRLTALVLSAVASTPFSNVLLVL 280

QY 339 HHANYHHNTDSLYFMVLIALLCLGSLNSCLDPFLYFVMS 377

DB 281 HYSNPSPEANGNLGYAVPSLALSTLNSCVDPFIYYVS 319

RESULT 40

AAW54080

ID AAR54080 standard; Protein; 361 AA.

XX

AC AAR54080;

XX

03-FEB-1995 (first entry)

DT

DE Epstein Barr virus induced (EBI-2) polypeptide.

XX

XX Epstein Barr virus; EBV; induction; detection; diagnosis;

KW lymphocytes; antigen; growth; differentiation; mediator;

KW infectious mononucleosis.

XX

XX Homo sapiens.

XX

```
PH Key      Location/Qualifiers
FT Modified-site 8
FT Region      /note= "Potential N-linked glycosylation site."
FT Region      32..57
FT Region      /label= Hydrophobic region.
FT Region      78..95
FT Region      /label= Hydrophobic region.
FT Region      106..127
FT Region      /label= Hydrophobic region.
FT Region      126..134
FT Region      /note= "This sequences motif
FT          (S-[I/V]-D-R-(Y/F)-X-X-X) is highly
FT          conserved among a large number of G-protein
FT          coupled receptors."
FT Region      150..168
FT Region      /label= Hydrophobic region.
FT Region      193..215
FT Region      /label= Hydrophobic region.
FT Region      242..265
FT Region      /label= Hydrophobic region.
FT Region      288..312
FT Region      /label= Hydrophobic region.
XX
XX WO9412519-A.
XX
XX 09-JUN-1994.
XX
XX 08-OCT-1993; 93WO-US09636.
XX
XX 25-NOV-1992; 92US-0980518.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
XX Birkenbach M, Kieff E;
XX
XX WPI; 1994-200183/24.
XX
XX N-PSDB; AAQ64126.
XX
XX DNA coding for Epstein Barr Virus induced (EBI) polypeptide(s)
XX and antibodies to EBI1, 2 and 3 - useful for detecting EBV by
XX hybridisation or by immunoassay
XX
XX Claim 8; Page 58-60; 84pp; English.
XX
XX EBV infected B lymphocytes recapitulate features of antigen
XX stimulation in enlarging, increasing RNA synthesis, expressing
XX activation antigens and adhesion molecules, secreting Ig and
XX proliferating. Unlike antigen stimulated B lymphocytes, EBV
XX infected B lymphocytes continue to proliferate (in vitro) as
XX immortalised lymphoblastoid cell lines. Because of the similar
XX effects of EBV and antigen, EBV induced genes are likely to include
XX mediators of antigen induced B lymphocyte growth or differentiation.
XX
XX Sequence 361 AA;

Query Match      20.2%; Score 431.5; DB 15; Length 361;
Best Local Similarity 31.6%; Pred. No. 1.7e-37;
Matches 95; Conservative 64; Mismatches 107; Indels 35; Gaps 9;

QY 102 YLRSLSTQVIPAIIYLLFVGVPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 YAHHSTARIVMPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVISDILPTTA 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 161 LPFKIAYHLNGNNVFGVMCRITTVFYGNMYCAILLTCWGNRYLATAPHTYQKLP 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 LPTRIAAYAMGFDMRIGALCRITALVFYINTYAGVNFMTCLSIDRFIAVHPLRYNKK 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 221 KRFSLLMCGIWMVWVLMPLPFI--LKQEHVLVHSEITTCDDVDACESPSSFRFY 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 RIEHAKGVCIFWIIIVFPAQTPLLLNPMKQE-----AERITCMEYPNFEETKS--LPWI 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 278 FVSLAFFGFLIPFVIIIFCYTTLTHK-----LKSQRILWLGYIKAVLLILVIFTIC 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 197 LLGACFIVGVLPILIIICVSIQICKLFRKTAQNPLTEKSGVNNKALNTIILIVVFLC 256
-QY 329 FAPTNIILVIH-----HANY-----YYHNTD-SLYFMYLIALCLGSLNSCLDPLFYVMS 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 FTPYHVAITIOHMIKILRFSNFLECSORHSFQISLHF----TVCLMNFNCMDPFIYFFAC 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 K 378
Db 313 K 313
```

Search completed: June 24, 2003, 12:03:08
Job time : 55.9362 secs